Regulatory, institutional and technological change in investment research

Alistair Haig

Doctor of Philosophy

University of Edinburgh

2019

Declarations

This is to certify that that the work contained within has been composed by me and is entirely my own work. No part of this thesis has been submitted for any other degree or professional qualification.

I produced this thesis while appointed as Early Career Fellow at the University of Edinburgh Business School.

In chapter 3 I use the term ‘we’ instead of ‘I’. This chapter is my draft of a working paper co-authored with Professor William Rees. A version is available on SSRN. Chapter 3 is based on my own data collection, analysis and discussion and is my own draft. All other chapters are my own drafts and are “solo” papers. Throughout other chapters I employ the term ‘I’.

Alistair Haig

Acknowledgements

I am extremely grateful to by principal supervisor, Professor William Rees, for supervising my work. Professor Rees has provided the opportunity for me to use a range of methods, to engage in practice and to develop my writing skills. I also wish to thank Professor Arman Eshraghi and Professor Jens Hagendorff, who acted as additional supervisors and to Professor Jo Danbolt, Professor David Cooper and Professor Seth Armitage for academic input throughout the process. Many other faculty members have provided valuable feedback and opportunities to connect with academics at other organisations. Informal discussions with visitors presenting at the School's accounting and finance seminar series helped to give me a wider perspective.

The thesis addresses live practitioner problems using unique industry data. It would not have been possible without ongoing practitioner interaction. I am grateful to many specialists and entrepreneurs in the research marketplace for providing questions, feedback and especially for sharing lively debates. Special thanks go to Sanford Bragg (Integrity Research), Mike Carrodus (Substantive Research), Jeremy Davies (RSRCHXchange), Ruari Hamlin (Bloomberg), Neil Scarth (Frost Consulting), Steve Kelly (EuroIRP), George Littlejohn (Chartered Institute of Securities and Investments), Vicky Sanders (RSRCHXchange) and Roy Shackley (Australian Securities and Investments Commission.

Abstract

This thesis comprises four studies, each designed to form an independent contribution. The central theme is change in the provision of investment research. The study focuses on independent investment research firms.

The first study documents the mechanisms used to pay for investment research. Through interpretation of documents and event participation, I use qualitative analysis to explain the payment mechanisms used in recent decades and how these mechanisms have changed. A key finding is that since 2017 many investment-management firms have shifted from an opaque, reciprocal arrangement, which resembles a gift exchange economy, towards a neoclassical economic model.

Investment management and research firms are adapting to change in payment mechanisms at a time of falling asset management fees and commission rates. One response is to automate. While many firms are exploring machine learning, little evidence on the potential for such approaches exists in the public domain. The second and third studies in this thesis address this gap by evaluating the effectiveness of machine learning in the investment research function. Both studies compare a large, global sample of analyst and machine learning outputs.

The second study evaluates the relative effectiveness of analyst and machine learning valuations in predicting future returns. Analysts make unbiased valuations using a standardized discounted cash flow model. Although neither analyst nor machine learning valuations serve as viable predictors of subsequent returns, when used together they offer credible explanatory power which can be increased by adding a price momentum factor. Analyst and machine learning accuracy could be improved by placing greater emphasis on past returns.

The third study uses regression analysis to compare analyst and machine learning risk assessments using next quarter volatility as the outcome variable. Both assessments can benefit by incorporating information from the other, and both are more accurate in countries considered to have superior informational environments. Analysts seem to underestimate risk associated with their own "buy" recommendations, but no equivalent miscalibration is apparent in machine learning predictions. The results confirm both the value of analysts' research and considerable potential for machine learning in financial analysis.

The final study summarizes the informational environment in light of regulatory, institutional and technological change. Stock coverage by analysts provides a window on the level of information available to investors. Archival and case study analysis indicates that analysts are providing research on most large and mid-cap listed companies in the US and UK, i.e., stock coverage remains wide. Coverage also remains deep, with most companies receiving coverage from a similar number of analysts and few companies covered by only one analyst. The exception is that fewer companies are now covered by more than 20 analysts, indicating that some surplus research has disappeared.

Independent research has expanded over the past decade. Despite this, the tendency for analysts to provide optimistic recommendations persists. Independent analysts rarely contribute to archives and are redefining coverage models. Analyst forecast datasets assembled by established data vendors understate the wider selection of research which is now available, but only to those who can find and afford it. Taken together, there is little evidence to date of a diminished informational environment for equity investors.

CHAPTER 1

Introduction

Decades of research has produced extensive documentation of securities analysts' outputs. As a result, we have a detailed understanding of their earnings forecasts, recommendations, target prices and other outputs such as the content of analyst reports. Many studies use these outputs to infer the role of analysts in markets. To complement these, researchers from various disciplines employ qualitative methods such as interviews. Yet the large literature on analysts rarely touches on the economics of their own industry. Expanding our knowledge of the market mechanisms would therefore be a useful complement to the extant literature.

The environment for analysts has undergone significant change in recent years. Two particularly disruptive changes have become evident. The first is the economics of analyst research. Since the start of 2018 investment managers operating in Europe have been required to change the mechanisms for paying for research. In most cases research is now paid for by investment management firms rather than end investors. This change and its implications have induced press coverage and industry discussion but few scholarly papers. Some existing theory and empirical evidence may be rendered obsolete.

The second change is the use of machine learning in investment research. A flurry of industry reports has emerged (see for example Deloitte, 2019). Machine learning could be a compliment or substitute for analysts. It therefore presents both threats and opportunities. This second change is topical in many knowledge industries but, in the area of investment analysts, such change remains largely undocumented.

This thesis comprises four empirical studies which address these regulatory and technological changes and the associated institutional change. Each is designed as a separate research paper.

The first study – chapter 2 – documents how the procurement mechanism has changed from a reciprocal gift-exchange system towards a neoclassical economic model. Payment data has generally been unavailable. Through several years of industry engagement and analysis of archives, I illustrate the five exchange mechanisms which have been used in the past four decades. The market-based system adopted via MiFIDII is replacing the long-established exchange of research (gift) and brokerage commission (counter-gift). Participants are now choosing to treat research as a cost to their firm thus marking a significant change in practice.

While many firms are exploring machine learning, little evidence on the potential for such approaches exists in the public domain. The second and third studies in this thesis address this gap by evaluating the effectiveness of machine learning in the investment research function.

The second study – chapter 3 – evaluates the relative effectiveness of analyst and machine-learning valuations in predicting future returns. This is a demanding task. Analysts make unbiased valuations using a standardised discounted cash flow model. Although neither analyst nor ML valuations serve as viable predictors of subsequent returns, when used together they offer credible explanatory power which can be increased by adding a price momentum factor. Analyst and ML accuracy could be improved by placing greater emphasis on past returns.

The third study – chapter 4 – provides further insight by contrasting the accuracy of risk assessments made by analysts and a machine learning process. Prior studies indicate that analysts tend to make informative risk assessments after controlling for risk characteristics. We compare analyst and machine-learning risk assessments using next quarter volatility as the outcome variable. Both assessments can benefit by incorporating information from the other and both are more accurate in countries considered to have superior informational environments. Analysts seem to under-estimate risk associated with their own “buy” recommendations, but no equivalent mis-calibration is apparent in machine-learning predictions. The results confirm both the value of analysts’ research and considerable potential for machine learning in financial analysis.

Machine Learning is now being used to perform some of the tasks performed by financial analysts. We evaluate the relative effectiveness of valuations made by analysts and machine learning in US and non-US samples. The analyst valuations predict subsequent returns. Machine learning valuations are contrarian until combined with the analyst valuation and lagged returns. Rather than choosing between mind and machine, investors would be best to use both sources of valuation. Our results indicate that the machine learning used by the early adopter in our study could be adapted to expand stock coverage in global equity markets.

The final study – chapter 5 – provides an overview of the informational environment in light of regulatory, institutional and technological change. Stock coverage by analysts provides a window on the level of information available to investors. Although some surplus research has disappeared, archival and case study analysis indicates that stock coverage remains wide and deep. Independent analysts rarely contribute to archives and are redefining coverage models. Archival data understates the wider selection of research which is now available, but only to those who can find and afford it. Taken together, there is little evidence to date of a diminished informational environment for equity investors. An overall conclusion follows.

CHAPTER 2

From gift exchange to neoclassical economics: how investment managers pay for research

2.1 INTRODUCTION

Investment analysts play a key role in financial markets (Spence et al., 2019). Brokerage firms supply analyst research to buy-side firms such as investment management firms, pension funds and hedge funds. Their summary recommendations are incrementally informative (Womack, 1996), as are their target prices (Brav and Lehavy, 2003) and their reports (Asquith et al., 2005). Contextual information and the basis for the analyst's investment thesis are even more valuable (Imam and Spence, 2016). Despite confirmation that analysts perform a valuable role, a deep-seated perception remains that the research they produce is free (Feng et al, 2019). How can analyst research be valuable yet given away free of charge?

This paper addresses this question by undertaking a qualitative study of the system which connects fund managers to external investment research, a marketplace which has undergone significant regulatory, institutional and technological change since the turn of the century. Most recently, the procurement mechanism has changed from one based on social and relational exchange to one based on a neoclassical economic model. Through qualitative analysis of archives and event participation, the paper illustrates the five exchange mechanisms used. The market-based system imposed by the UK regulator, and adopted across Europe due to the Second Markets in Financial Instruments Directive (MiFID II), now performs the market globally.

Invoices and budgets display explicit prices which replace the exchange of research (gift) and brokerage commission (counter-gift). Participants are now choosing to treat research as a cost to their firm, thus marking a significant change in practice. Investment managers now benefit from an expanded choice of research suppliers and better value for money but can no longer use research without first negotiating payment.

The analysis presented, which is based on event participation and analysis of documents, indicates that research is not free. Brokers offer research with the objective of obtaining payments which can be disproportionately large. For successful brokerage firms, and the analysts who work there, this results in substantial asymmetric payoffs at the expense of other brokers who can lose out completely on the commissions paid to trade any given share. The opaque nature of this reciprocal system is beneficial both to producers and consumers of research, not least because the cost of research is charged to the end investor.

Perhaps such complexity provided a smoke-screen for analysts operating in tandem with dealmakers and thus creating conflicts of interest (Mehran and Stulz, 2007). Analysts had come under little scrutiny during the late 1990s technology-stock boom but media coverage of the subsequent market crash revealed that investment bank analysts had misled investors with overly positive forecasts and recommendations. Making deals for corporate clients had disadvantaged all but the largest investors. Analyst conflicts had become headline news in the financial press and litigation against analysts proceeded (Wu et al. 2017).

Eventually this led to investigations led by the US State Attorney and to US congressional hearings into biased recommendations by some Wall Street analysts, notably, buy recommendations made to attract interest in companies issued by the investment bank, sometimes accompanied by warning stock selecting investment management clients. The investigations resulted in the 2003 Global Analyst Research Settlement (hereafter GS) through which ten investment banks were required to pay a total of $1.4bn in fines and to subsidise independent research, accompanied by tightening of New York Stock Exchange and other US stock exchange rules. Taken together, these rules marked out clearer separation between analysts who serve investment management clients and investment bankers who are paid by the corporates covered by the analysts. For example, analysts could no longer report to investment bankers and the publication process mitigated interference from deal makers. The GS had global ramifications as the firms involved had substantial non-US operations. Many non-US firms and regulators adopted similar practices.

The European regulatory environment also goes back several decades. The UK regulator has been the main driver of policy on analyst research with other EU regulators taking a more relaxed approach towards this topic. Coincident with US congressional hearings in the early 2000s, the UK government commissioned a report into the UK investment management industry by the former chairman of a large UK investment management firm (Myners, 2001). In the report Lord Myners, an industry veteran, recommended a very different approach to that taken by US authorities: the emphasis was not on rules for brokers but instead on incentives for investment managers. Myners recommended that dealing commissions should not be used to pay for research and instead should be separated or 'unbundled' in order to create a more efficient market:

> *Clients' interests would be better served if they required fund managers to absorb the cost of any commissions paid, treating these commissions as a cost of the business of fund management, as they surely are… Under this system, the incentives would be different… Fund managers would choose which services to buy and which to provide themselves…The pressure would be to purchase only those services which contributed to such returns, and to do so in the way which is most efficient.*
>
> (Myners, 2001, p11-12)

In 2003 the UK regulator (then known as the FSA) engaged in an industry consultation based on Myners' recommendation. The outcome of this consultation process led to more moderate innovation. In the UK regulator's own words: The FSA initially proposed more radical reform:

> *…the responses to the consultation argued that alternative approaches could deliver similar improvements at less cost and impact to the industry. The FSA was therefore persuaded to work with the grain of an industry-led solution and evolving market practices* (FCA, 2014, p15).

The result was the FSA 'use of dealing commissions rules' which came into effect in 2006, requiring non-execution commission to be limited to research and not used for other purposes, thus differentiating it from the concept of soft dollars, where commissions charged to clients could be used to pay for data vendors and other business services, meaning the end investor pays the bill. It also encouraged the use of Commission Sharing Arrangements (CSAs). The Centre for the Study of Financial Innovation (CSFI) commented as follows:

> *The system also falls short of the vision of full unbundling envisaged in the initial FSA proposals in the wake of the Myners report (FSA CP 176, April 2003), under which fund managers would have been required to pay for research out of their own fees, rather than from clients' commissions. Instead, the … rules required only that fund managers provide adequate disclosure of these costs to their clients under industry-led guidelines developed by the Investment Management Association and the National Association of Pension Funds.*
>
> CSFI (2011) p22

The FCA concluded, based on consultation and evidence presented by an economic consultancy (Oxera, 2009) that the rules had resulted in a meaningful improvement in the alignment of incentives and the efficiency of the market. Again the yardstick was market efficiency. This status quo remained in place for several more years. The consensus among those surveyed (in 2010/11) was that the prospects for further action by the regulator were not high. Having commissioned Myners' report, acted on its findings and declared itself satisfied with the implementation of the resulting rules, many felt that the FSA's focus would lie elsewhere. In the wake of the crisis, the regulator was instead more concerned with issues of financial stability. (CSFI, 2011, p18) Given the regulator's comfort in Oxera's evaluation, it may be that the fund management industry believed that partial compliance had been accepted as sufficient. Many organisations were focused on surviving the financial crisis and the FSA itself had many other areas of concern. The CSA payment device was emulated in the US where the regulator (SEC) introduced CSAs (referred to in the US as Client Commission Agreements, or CCAs, which are largely equivalent to CSAs) in 2006. Although these arrived almost concurrently with UK CSAs, the range of research providers who would accept third party payments was limited. This legal point impeded the US application of the new market mechanism until 2010 and change was therefore initially constrained (Frost Consulting 2014).

MiFID II, which came into effect in January 2018, requires investment management firms to separate payment for research services from execution and to negotiate terms in advance. They can only pass on the cost of research to clients if they develop ex-ante budgets and pay via a strictly regulated Research Payment Account (RPA); the alternative is to charge an expense to their firm's own profit and loss account.

The new European legislation imposes more stringent rules on the payment, procurement and provision of investment research. The directive requires investment management firms to separate or 'unbundle' payments for external investment research from commissions paid for trade execution. The new rules seek to prevent brokers from using research to encourage investment managers to trade and therefore impose unnecessary costs on end investors. Investment managers are now required to absorb research costs from their own profit and loss account, alongside their other business expenses or to pass on the costs to clients under strict rules via a Research Payment Account (RPA). The rules include pre-arranged budgets to control the overall costs and transparent information on the payments made to each research provider. Investment managers must document an explanation for how each research service enables the firm to make better investment decisions. Accounting and administration requirements are therefore significantly more onerous post MiFIDII, even in the UK where the regulator had tended to be more demanding than most regarding research payment rules.

By treating research which is not paid for as an inducement, MiFIDII changes the procurement process. Investment managers must negotiate prices with research providers in advance and must not use research that is not paid for. This last point is particularly important as buy-side firms need to ensure that their staff do not receive research which has not been explicitly purchased. Research providers must also desist from attempting to supply research prior to payment.

MiFIDII applies to investment managers and investment providers who do business in any of the 31 European Economic Area (EEA) countries. The impact is therefore global as many non-EEA firms have a presence in the EEA. Additionally, the EU directive has important and ongoing ramifications for US regulators. This is unusual as it is typical for influence to flow from US regulation into the rulebooks for other nations and into compliance manuals for global investment firms. In 2017 the SEC has implemented temporary regulations to help firms comply with MiFIDII; in November 2019 these were extended to 2023 (SEC, 2019). The US regulator has announced ongoing work aiming at a more permanent solution.

Feng et al. (2019) also report that US institutional investors are collectively urging US regulators to implement MiFIDII-type rules around research payment on a permanent basis. Some European end investors have reduced the overall cost of investing as a result. If US end investors are paying more this implies that they might be subsiding non-US investors, a disparity which will not sit well with the fiduciary duty placed on institutional investors, especially some of the best-resourced pension and sovereign wealth funds in countries such as US, Canada, Australia and Singapore. Such a call to action was predicted in Haig and Scarth (2017).

My key finding is that the procurement mechanism has changed from one based on social and relational exchange to one based on a neoclassical economic model. I depict five exchange mechanisms which have been used in the past four decades. The market-based system adopted via MiFID II is replacing the long-established exchange of research (gift) and brokerage commission (counter-gift). Participants are now choosing to treat research as a cost to their firm, thus marking a significant change in practice.

2.2 PRIOR RESEARCH

Barker (1998), Healy and Palepu (2003) and Bradshaw (2009) each depict information flows connecting analysts, companies and investors but none show explicitly who pays for research or how that payment is made. Despite calls for the direct study of analysts (Bradshaw, 2011), we know very little about how these flows are paid for. Few scholars have investigated the payment mechanisms and, even in practice, detailed knowledge of the marketplace was considered a specific domain rather than essential for most investment professionals.

This gap seems odd given that we would expect information flows to be important to investors, not least because the scale of the market for analyst research was large enough to be cited in reports such as that of Myners (2001). Even if the research marketplace has seldom been investigated, prior studies have established that investment research is valued by investors. Barker (1998) ranks the perceived value of sell-side analysts' outputs and finds advice, such as contextual information and access to management, to be more valuable than predictions. Extensive surveys of research on analyst forecasts by Ramnath et al. (2008) and Kothari et al. (2015) do not document research procurement, nor do surveys of analyst practices such as Block (1999), Imam et al. (2008), Pinto et al. (2015) and Brown et al. (2015, 2016). Participants in this market have expertise in valuing companies; we will see that many seem not to have applied these skills to valuing or pricing analyst research itself.

Additionally, equity valuations, earnings forecasts and trading recommendations have been consistently found to be of secondary importance to analysts' investment management clients. Brown et al. (2015, 2016) echo prior research in finding that analyst access to management is prized by buy-side clients, although UK brokers have been prevented from charging for direct access to companies since 2005, a point emphatically restated in a 2014 FCA clarification. Brown et al. (2016) underline the importance of advice relative to predictive accuracy; Imam and Spence (2016) emphasize that much work in the mainstream accounting and finance journals is based on the flawed presumption that forecasting accuracy is the main objective of analysts. Spence et al. (2019) reinforce the importance of advice over summary outputs such as earnings forecasts, target prices and investment recommendations (Barker, 1998; Beunza and Garud, 2007). Bradshaw et al. (2017) gather together the literature on the role analysts take in supplying information to the marketplace.

Imam and Spence (2016) follow Beunza and Garud's (2007) classification of papers into three strands: "*mainstream accounting and finance; neo-institutional and behavioural work; and emerging sociological perspectives on financial intermediaries, of which analysts are one group*" (Imam and Spence, 2016, p228). While the breadth of this review is particularly valuable to our understanding of the role of analysts, it confirms that prior studies have not addressed the marketplace for their products or services.

2.3 THEORETICAL FRAMEWORK

The relationships between analysts and their investment management clients have been shown to be crucial (see, for example, Barker, 1998; Brown et al., 2015; Spence et al., 2018). We can therefore expect to find relational ties built into the exchange mechanism (Granovetter, 1985). It would be useful, however, to have a theoretical framework to help us understand the nature of these embedded relationships, to explore how they work and why they exist.

2.3.1 Gift exchange systems

Research services range widely as does their value to each investor. For example, a meeting with an analyst to discuss catalysts for a neglected value stock could be of no interest to a growth investor; a perceptive inflation report might guide a macro-driven investor but be irrelevant to another; quant screening could assist a stock-picker but might not suit a thematic process. Research is made up of a complex array of services: calls, meetings, tours, models and a wide array of written material extending far beyond the "single name" research which features in prior studies. Gift exchange has been found to be effective in such marketplaces (Offer, 1997). The study of gifts has, initiated by Mauss in 1925 and republished at the start of this century (Mauss, 2002), has become an important theme in economic sociology. Mauss (2002) contested the belief that barter predated monetary systems of exchange. Instead of barter, gift exchange was the norm: a gift is given, received and acknowledged with a counter-gift. Participants in an economy develop expectations of the size and nature of the appropriate counter-gift; there may also be a hierarchy where gifts are offered to members of society who are recognized to have more power. Tribes and their members would therefore not consider personal interests, property or freedom; much of the assumptions behind a neoclassical model of exchange would seem irrelevant. Counter-gifts persist even when fiat money has become well established, e.g. a birthday gift might be offered with some expectation that one's own birthday is remembered. Important aspects of gift exchange can therefore be found in modern market systems.

Gift exchange is not barter; rather it is an holistic means of dealing with individual transactions; rather, it marks out a system to exchange not only goods but also services such as banquets, and military assistance. Even tribe members could be exchanged. The exchange has economic, cultural and religious meaning. In this system, the size, type and magnitude of a counter-gift is well known to tribal members. Failure to reciprocate is believed to bring misfortune or may even instigate attack by the giver. Mauss points out that these characteristics are also found in history, from India to Greece, and that it is not coincidental that the Greek words for “poison” and “gift” have the same root (Mauss, 2002).

Akerlof (1982) applies Mauss’s theory to show that workers adapt their behaviour according to the apparent generosity of their employers. Workers respond to extraordinarily high (low) wage levels by exhibiting high (low) productivity. Bell (1991) generalizes this using formal economic analysis to show that a neoclassical market model is a special case of gift exchange. In another attempt to incorporate sociological conceptions of exchange into a neoclassical model, Offer (1997) finds that gift exchange is most effective where goods and services are unique in nature, costly to produce and have multiple dimensions of quality. In these situations, personal interactions are valuable. Neoclassical models are likely to be less efficient. We can therefore expect to find reciprocation not just in the tribal societies studied by anthropologists and historians but also in established market economies. Gift exchange theory provides a lens to examine the research marketplace prior to the arrival of invoices, contracts and prices. In the following section I employ a second theory to study the introduction of a neoclassical market model.

2.3.2 Performativity

The second key area of theory which I pursue in this chapter is the study of performativity which has developed in new economic sociology, and related disciplines, known collectively as social studies in finance. MacKenzie et al. (2007) define generic performativity as the use of an aspect of economics, such as a theory, model, concept or procedure, by participants in the economy. For performativity to be effective, the practical application of a given theory must have some effect on economic processes. Two special cases of effective performativity exist. The first, ‘Barnesian’ performativity, is the subset where the “effect is to alter economic processes to make them more like their depiction in economics” (MacKenzie et al., 2007, p. 67). The second special case, counter-performativity, exists where economic processes become less like their depiction in economics. MacKenzie (2008) shows that option pricing theory, widely considered to be a key achievement of 20th-century economics, was Barnesian performative, and also sometimes counter-performative, in the exchange markets for financial options. Option theory conferred legitimacy on the option markets, allowed contracts to be accurately valued and entered the parlance of traders.

The edited work of MacKenzie et al. (2007) collects key debates on the topic. In this volume, Michel Callon develops ideas from his earlier work (Callon, 1998) to argue that our understanding of markets could be improved by taking into account the features of the marketplace. These features can be significant and institutional details can improve economists' analysis. In MacKenzie et al. (2007), Callon (2007) argues that markets, such as those following rules of neoclassical economics, are the norm, which we could think of as being within a frame; yet exchange is impacted by social relations which surround this frame. In the same volume, Miller (2007) argues for the exact opposite; using evidence from an economy in rural India, he argues that social connections form the essential mode of exchange, i.e., are within the frame, and are complemented by markets. Consumers only a need to go to market when there is a shortage; producers need only supply their excess. While a rural village seems far from the workplaces of investment professionals, Miller's argument holds some plausibility in our own daily life and in corporate affairs, many of our exchanges are based on relations first and the market second.

Another aspect of the present study is that the UK regulator has set out to design a neoclassical market to improve the outcomes of investors. Garcia (2007) shows how perfect competition is used as a model to construct a physical auction market. The setting, a network of co-operative growers in a French region not known for its strawberries, seems far from the modern stock exchange, yet we must remember that investment analysts constitute a production market even if their work immerses them in exchange markets. Garcia's fieldwork depicts a traditional agricultural region where distribution is arranged through a combination of personal relationships and brokers. This traditional arrangement discouraged specialist production of strawberries, due to suppliers' expectations and delays in payment. In the early 1980s, the region's local government set out to promote the region's strawberries. As part of this initiative, they sent a young civil servant to act as a consultant. After several unsuccessful attempts to modify the behaviours of existing co-operative members, the consultant applied theory learned from undergraduate economics courses. Together they used a model of perfect competition to construct an auction market where buyers and sellers could not see each other and with both payment and delivery taking place during market hours in purpose-built facilities. The consultant and at least the better-informed market participants would have been aware of the opening of a new auction market in another province which may also have prompted the decision to adopt this approach. A minority of relatively knowledgeable growers embraced the consultant's proposal and the market was performed (Callon, 1998) according to the textbook image of perfect competition. The social construction of a neoclassical market in early 1980s rural France is supported by participants in preference to other systems. The research market differs in that participants have largely resisted moving from relational exchange to a neoclassical market: the textbook model is imposed on the investment research marketplace.

2.4 METHOD

The objective of this study is to explain the market mechanisms used to pay for investment research. In 2013, research procurement was considered a specialized domain even within practitioner circles. Most investment managers relied predominantly on brokers for external research. They understood the principle that research commissions are paid by the fund, i.e., by end investors, rather than from their employer’s own money, but few were up to date with the letter of the regulatory requirements in this area; instead, they would rely on dealers or compliance specialists. Some senior investment managers and executives had become well versed in the processes.

Participation at events is useful in the study of a complex and dynamic setting. The events listed in Table 2.1 usually featured panel sessions through which I could follow the debate as the regulations changed.

2.4.1 Participation in industry events

Event presentations and question-and-answer sessions provided perspectives from senior professionals representing investment firms, regulators and industry bodies. These events ranged from small, specialized round table discussions to conferences with up to 300 delegates. I took the opportunity to ask questions both during proceedings and on a one-to-one basis with speakers and participants. Participation also served as further reassurance that the project was of relevance to stakeholders. Table 2.1 provides a schedule of events attended.

Event proceedings were developed from fieldwork notebooks and combined with transcripts or summaries prepared by the event organizers. Chatham House rules often prevailed, where direct quotes cannot be made without permission, thus encouraging candid discourse. Event participation aided identification of relevant reports by investment firms, regulators, consultancy firms and the financial media. Events also allowed for informal discussions and I would often follow up with telephone calls or subsequent face-to-face meetings.

2.4.2 Documents and secondary analysis of questionnaires

The first major proposal to reform the economics of investment research was that of Myners (2001). In a report on the UK equity market, commissioned by the UK government, the former chairman of a large UK investment management firm called for the use of dealing commissions to be banned. The UK regulator responded with a series of consultation and discussion papers which were often mirrored by industry consultation. The documents consulted are listed in Figure 2.1. These archives are recent, and this provides evidence in itself that the market for research was not at the forefront until the mid-2000s or later. Taken together with events, these documents served to establish the historical timeline of regulatory actions as they represent collective views and are based on their own primary research either through surveys, regulatory supervision or consultation. The documents were categorized by the type of organisation authoring or sponsoring the report. Buy-side and sell-side market participants verified the credibility of all sources.

To understand the introduction of a competitive marketplace for investment research, I use an interpretive approach, working back and forwards between the available data, which expanded considerably during the study, and the economic sociology literature on production markets. The epistemological approach is designed to maximize understanding. The data is collected and analysed in order to describe and explain the changing marketplace. The model developed by Barker (1998) served as a starting point. In 2014, two papers became available: a summary of how selected research providers are responding to a decline in broker commissions (Healy, 2014) and a study of the broker votes collected by a single brokerage firm in the mid-2000s (Maber et al., 2014). Neither examined the importance or changing nature of payment mechanisms. Haig and Rees (2016), however, add payment flows to Barker's model (Barker, 1998); I have expanded this work in the diagrams presented in this chapter.

Each document was used to identify features of each mechanism which correspond to neoclassical market theory or to relational exchange theory. Five exchange mechanisms were identified and a timeline created to chart changes in the mechanism employed in the UK. Event participation facilitated understanding of the mechanisms and the extent of adoption. Events were also helpful in plotting a timeline which is summarised in table 2.3.

2.5 FIVE WAYS OF PAYING FOR RESEARCH

2.5.1 Mechanism 1: Paying on execution (UK prior to 2000)

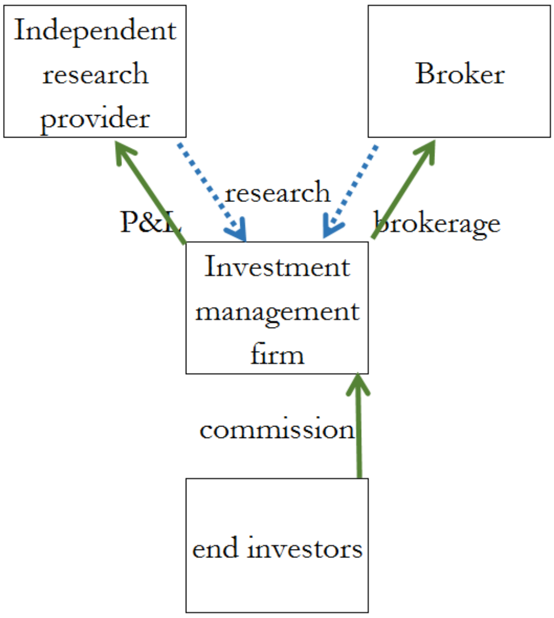
The practice of allocating commissions based on broker research was commonplace by the 1930s and its history is neatly summarized by Bradshaw et al. (2017). Until the deregulation of major stock markets, such as the New York and London stock exchanges, in the 1980s, the commission rates payable to brokers were fixed by law and so there were limited means for brokers to compete for the custom of investors. A broker with renowned analysts could attract business and influence buy or sell decisions through the distribution of research reports, which also formed a basis for salespeople to call clients in the hope that they might trade through the firm:

> (*i)nvestment research has traditionally been seen as a cost centre for brokers, often un-priced and given away for free in the hope [that] trade ideas will generate trading commissions for the research provider* (Quinlan, 2017, p. 4).

Fund managers access research without contracts or even prices in this social and relational system. Brokerage firms with the best analysts can attract more trades and therefore have the highest commission revenue. We can think of research as a gift offered by brokers to (many) fund managers, one which is received when investment managers use the research. The counter-gift takes the form of a commission which, since it is calculated as a percentage with no upper limit, can be large. Broker research is designed to incentivize fund managers to trade. Brokerage analysts need to balance this tendency to issue over-optimistic recommendations with the need to develop a reputation (Jackson, 2005).

Suppose a brokerage analyst has just released her latest report, a “buy idea” on the company, in which her recommendation has moved from “Hold” to “Buy”. A fund manager incorporates some of the new report into his own analysis and calls the analyst to discuss some assumptions and details (note that he has a duty to have reasonable basis). This process leads him to purchase shares in the subject of the report and he instructs his dealer to purchase the stock from the broker in recognition. The fund manager has accepted the research as a gift and uses brokerage commission to make a counter-gift. Note that single-name research is rewarded by trading shares in the named company.

Figure 2a: On execution and broker vote payment mechanism



Notes

The flow of information (research) is shown in dotted lines. Payment flows are in solid lines.

Panel A shows the flows for on execution and broker vote mechanisms, the typical methods used prior to MiFIDII. Brokers supply research to investment managers. If the research is useful to investment managers the latter reciprocate with dealing commissions which bundle together payement for execution and research. The payment for research and dealing is bundled together and deducted from the value of the end investor's fund.

The broker vote uses the same mechanism as on execution. The broker vote is the internal process used to arrange payments each investment management firm. Rather than simply making payments each time a trade is placed with a broker, the investment firm attempts to pay brokers proportionally based on the quality of their research.

Additionally, investment managers can pay independent research providers but this can only be done by incurring an expense. In practice, most research is procured from brokers using dealing commission because this reduces the costs incurred by the investment management firm.

Following industry norms, there is no contract specifying the amount to be paid and so the broker wins all the research commission. The broker will only be paid when a trade is allocated. Brokerage firms are playing for a ‘winner takes all’ outcome with each fund manager. No payment is made to analysts at other brokerage firms who also cover the stock.

Despite being freely available to clients, written research frequently remained unread. Not all fund managers use the research and there may be a number of reasons for this. They may not be interested in the sector, they may already hold a large long position in the company, or they may prefer other analysts who cover this sector. Survey evidence (RSRCHXchange, 2017) shows that although fund managers read written research regularly and find it to be useful, a fund manager can be expected to read up to 5% of the reports made available to him. Many competitors will leave the report unopened; the automated email is deleted, archived or just ignored. The gift is not acknowledged.

During the year, analysts provide research to a list of investment management firms. They send research reports by email and, aided by sales colleagues, attempt to call and meet clients who are in a position to allocate brokerage commissions. Some gifts are acknowledged: clients reply to discuss stocks in her sector either by phone or in a meeting; others simply read and consider her report. Clients who value the “idea” may reciprocate with a counter-gift in the form of brokerage commission. Counter-gifts range in size because research commissions are calculated on an ad valorem basis. The manager of a smaller fund will make a smaller payment. It is quite common for the price paid for research to be different, primarily because of the size of the trade, a function of fund size, market levels and fund turnover.

It is also possible that some buy-siders read and use an analyst's research, but choose not to pay. These firms may view research as part of a bundle of services provided by each broker, or an advertisement to encourage trading. The industry is divided into participants who are keen to identify and pay for research which they believe will improve their investment performance and those who view it as a form of advertising. Not all of those who receive the research acknowledge receipt by paying commission. The latter may find themselves excluded from research services in future (most likely the restrictions would apply to analyst time rather than written research) unless they can demonstrate payment: in a gift exchange system, *"(t)he penalty for failure is exclusion*" (Offer, 1997, p. 453).

Reciprocation takes place not just at the firm level but also at the individual level. Hospitality, such as a good lunch, may prompt a fund manager to consider trading; he may feel a personal need to reciprocate by paying commission, even though the counter-gift is funded by the end investors rather than his firm. It is also possible that the lunch facilitated cognition by extending his knowledge of a firm or sector. In the latter case, the end investor may have benefited due to the investment managers' learning process. Gift exchange at both the personal and organizational level become entwined. Brokerage analysts therefore seek cordial relationships with fund managers to encourage reciprocation (Cialdini, 1993).

Traditionally only the executing broker could be paid for research and brokers competed for commissions on the strength of their analyst research. The analyst with the second-best research on the same company is not rewarded at all. Asset management firms can only pay for research by instructing trades as shown in Figure 2a While the percentage paid for execution and research can be specified (since 2007 this is essential in Europe due to MiFID II requirements for 'best execution'), the payments are said to be bundled. An investment management firm needed execution arrangements, specifying commission rates but not research quantities, to be in place with each broker to pay for research. There exists a one-to-one mapping between research and execution. Fund managers often maintain 20 broker arrangements, rather more dealing counterparties than the buy-side dealers deem necessary for executing trades, in order to tap into a broader selection of research.

A textbook model of perfect competition requires that there are many buyers and sellers (atomicity) who take prices as given and can see the supply and demand of products (transparency), standardized products (homogeneity), and freedom for buyers and sellers to enter or exit the market. Paying for research on execution is an opaque, relational arrangement conferring benefits on both producers and consumers of research; it stretches the assumptions of perfect competition. These benefits reflect a theme throughout the event participation and document analysis and help to explain the general resistance to a transaction-based mechanism.

2.5.2 Mechanism 2: Broker vote (UK 1990s to 2017)

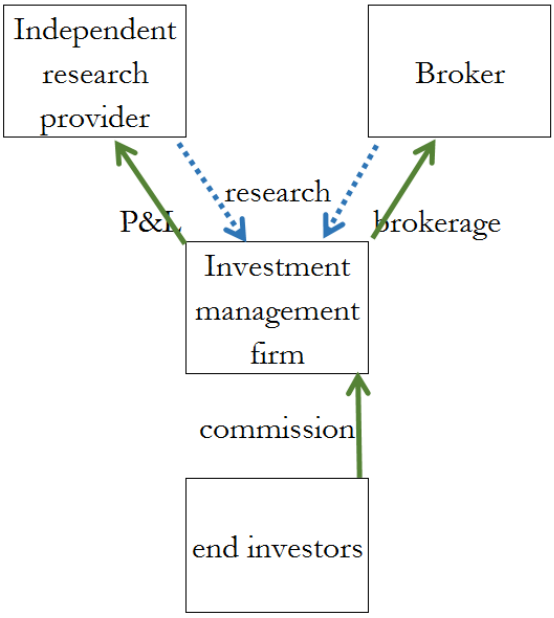
In a broker vote system, fund managers within a firm rank (vote for) brokerage analysts based on the value of research provided in a given period. The fund management firm then allocates commissions in proportion to the rankings. Although the parties can calculate the percentage share of the commission pot, no invoicing exists, and it is unusual to find prices expressed in currency units. The absence of explicit prices casts this arrangement in sharp contrast to execution commissions which must be carefully measured to minimize the cost borne by clients (the adverse effect on fund returns caused by deduction of commissions) as part of best execution as mandated by the original version of MiFID which took effect in 2007.

Research commissions, aggregated within the buy-side firm and referred to as the commission pot, vary from month to month. The size of the commission pot depends on several factors which are time varying. The first is portfolio turnover: buying and selling shares generates commission. The second, assets under management, in turn depends on the value of the shares in the fund, and cash flows, such as investments and redemptions, made by end investors.

Commissions are awarded to brokers by the buy-side firm based on feedback from dealers, for efficient execution of trades, and investment managers, to pay for research. Investment managers within a firm might each rank three analysts per sector. This information is then aggregated to form the broker vote. The process is also designed to capture not only recommendations but also the wider range of research services, including strategy, economics, customized projects and quantitative analysis (Maber et al., 2014). These services are not easily accounted for when paying on execution.

The vote is recorded in a spreadsheet and then presented at a broker review meeting. The buy-side dealing desk then endeavours to allocate trades to brokerage firms according to the votes in the coming period, e.g., calendar quarter. Note also that actual payment takes place at the time of the trade, which might be some months after the broker vote or review meeting. There is no means for research commissions to be accrued beyond the period for this vote. For example, suppose there is a year in which a fund manager is busy buying shares because investors are placing fresh money into his fund. The commission pot will be large and can be used to pay for an abundant supply of research. The same investment manager might then trade very little in the following year and might afford only limited research. This direct linkage between research and dealing provides fund managers with an incentive to trade in order to buy research rather than to meet portfolio objectives. Consequently, the system may be inefficient for the end investors who bear the cost of transactions.

Figure 2b: Broker vote payment mechanism



Notes The flow of information (research) is shown in dotted lines. Payment flows are in solid lines. The broker vote uses the same mechanism as on execution. The broker vote is the internal process used to arrange payments each investment management firm. Rather than simply making payments each time a trade is placed with a broker, the investment firm attempts to pay brokers proportionally based on the quality of their research. Additionally, investment managers can pay independent research providers but this can only be done by incurring an expense. In practice, most research is procured from brokers using dealing commission because this reduces the costs incurred by the investment management firm.

While the arrangement adds some structure and formality, no contract is in place. The process is more organized, but it is important to note the amount paid is still expressed as a percentage of the ‘commission pot’, i.e. the total research commission paid in a single quarter or year. The actual value of shares traded could turn out to be much larger or smaller, and different payment amounts could be made in different periods for a similar research service. One investment management firm could multiply its assets under management in a single year; another could fail and close for business.

Compared to trading on execution, the broker vote conveys more information about what research has been accepted in the past period, and might be extrapolated to indicate what research is expected to be accepted in future. By citing examples of useful service, a fund manager can signal information on demand: not the demand curve of an economics textbook but rather a demand schedule (White, 1981) and therefore obtains some basic management information which can be used to compare payments made to different brokers and analysts. In turn, this could inform negotiations and demonstrate to clients that there is a process for commission payment. The investment manager can compare the payments made to different brokers and even the analysts within each brokerage.

The lack of price information does, however, lead to a number of limitations. While buy-siders can identify the percentage of the commission pot allocated to each broker, it may be difficult to refine this to determine exactly how much is paid to each analyst. It would be difficult to establish how much was paid for different types of service such as reports, meetings or customized projects. Even if this is possible, the actual price is not known until next year's dealing is performed. Neither the buyer (fund manager) nor the seller (broker) have a way to price research on an ex-ante basis; both can make a calculation of the price if they wish, but with no regulatory requirement or client pressure there are likely to be other more pressing matters to deal with. The accountant seeking information about the value of research services would find no invoices or contracts; instead, they will only find an empty box.

Three adverse consequences result from these limitations (CFA, 2014). First, because the vote arrives at percentages, the price of a certain service in dollar terms can fluctuate from year to year due to changes in funds under management. An investment manager requiring exactly the same research would be charged more, and the total payment is entirely determined by changes in stock prices or fund inflows. Second, the investment manager needs to trade in order to pay commissions to the broker, which creates the incentive to trade even if transactions are not required. Third, broker votes have often failed to provide useful feedback to brokers regarding the services required. For these reasons, the UK regulator pronounced that the broker vote was 'inherently flawed' (FCA, 2014). A large survey of UK sell-side firms revealed that the broker vote system provides feedback which is lacking in detail, accuracy and timeliness (Extel, 2011, Figure 2.1). As a result, fund managers may well have breached their fiduciary duty to act in their clients' best interest.

The broker has some incentives to maintain this form of exchange. If a strong relationship can be established, then there is a greater chance of commission payments in future. If stock markets increase in value, or if the buy-side firm attracts clients, brokers will earn higher commissions due to the percentage system: an asymmetric, option-like payoff with very large upside and limited downside (zero payment).

The investment management group might also benefit from this arrangement. The buy-side firms, and the fund managers who use research to inform decisions which will define their career success, continue to benefit from research at the expense of their end investor clients. The process allows research to be consumed before deciding how much, if anything, should be paid. The fund manager can request access to research then decide how much to pay. Fund managers can find themselves in a position of power (Imam and Spence, 2016).

The broker vote introduces some formality to the ritual by determining the frequency of exchange and appropriate size of counter-gift to each broker. If research commission is not forthcoming, the fund manager is likely to be approached by the broker to press for commission to be paid out of trades in the near future, reminding the asset management group that the counter-gift is expected. Ultimately, if commission is not allocated according to the vote, the broker may choose to suspend provision of research services, e.g. by ceasing to provide analyst time, or at the extreme by stopping sending written research. The latter approach would be quite uncommon as it reduces the potential for incoming counter-gifts in future.

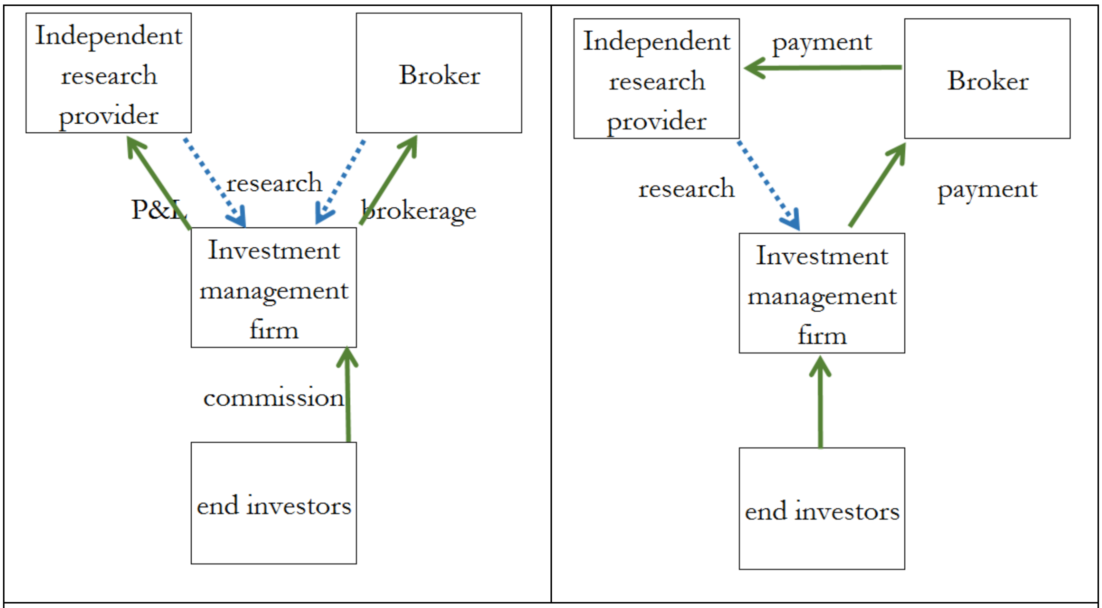
The dominance of a social and relational mechanism over one aimed at revealing prices also stands in sharp contrast to the stock markets which form the topic and motivation for most investment analysis. Fund managers and analysts exist by virtue of competitive markets with transparent prices.

In summary, the broker vote regularizes the ritual of exchange. It improves orderliness but maintains an opaque, relational system. It facilitates negotiation of payments made by a single fund management firm to multiple brokers for analyst services. It makes the time interval between provision of research and payment more explicit. It introduces order and helps to organize the payment to brokers from the commission pot. The vote is, however, only a crude form of accounting based on percentages. There are no contracts, invoices or explicit prices. Votes do not remove the incentives for brokers to produce a large volume of research or for fund managers to trade more and pay more than they would in a neoclassical economic model. The mechanism itself does not allow the fund manager access to research which was previously unavailable.

2.5.3 Mechanism 3: Commission sharing agreements (UK 2005–2017)

Prior to the introduction of Commission Sharing Agreements (CSAs), research and execution commissions were bundled together whenever an investment management firm made trades. In the mid 2000s, regulators in the UK, US and some other markets created a mechanism allowing buy-side firms to place research commissions in an account from which they can pay any research provider who will accept a CSA payment; this includes not only brokers but also independent research providers. The broker who executes the trade earns all of the execution component, but the research component can be redirected to one or more other parties as shown in Figure 2c. For example, trade execution could be provided by a given broker with the research component paid away to other parties to pay for provide research. CSAs therefore enable the “unbundling” of execution and research components. Indirectly they also reveal prices because payments from the CSA account require an invoice showing the price to be paid in currency units rather than as a percentage of the commission pot. No ex-ante budgeting is required. Since CSAs are voluntary, it is common to find a hybrid approach which utilizes CSAs as part of a traditional broker vote.

Figure 2c: CSA Payment Mechanisms



Notes

The flow of information (research) is shown in dotted lines. Payment flows are in solid lines.

Left panel shows the mechanism underlying both execution and broker vote mechanisms which we have seen in the previous two sections.

A CSA (right panel) allows an investment manager to use a given broker for execution and make payments to a separate account to pay for research. The investment manager retains discretion as to how the CSA account balance is used to pay for research from other research providers, including independent (non-brokerage) firms. The process allows the investment manager to maintain fewer brokerage counterparties while accessing a wider range of research suppliers. CSAs require invoices to be issued this placing a monetary value on each research service. This contrasts with broker vote allocations which are expressed in percentage terms and therefore vary through time with the investment management firm's assets under management.

CSAs became available in 2006 in the UK, US and some other markets. CSAs were optional and, although initial adoption was slow, by 2015 many buy-side firms used a combination of CSA and broker vote allocations.

The CSA system contrasts with the broker vote in a number of ways. First, the range of available research becomes wider. Fund managers can use CSAs to pay for research from non-execution counterparties, bolstering independent research firms of various shapes and sizes. Second, invoices specify prices in finite currency values rather than percentage votes. CSAs make prices explicit in invoices which are formatted just like any other commercial invoice. Third, research need no longer be purchased at/near the time of the trade. Fourth, the invoices require monetary values, and the cost of research becomes less dependent on the size of the fund.

A key driver for CSAs was the UK regulator's competition logic. Having motioned a ban on research commissions in 2003, the UK regulator chose instead to provide the option of a cost-based system. In the words of the leading UK-based consultancy in this area this allowed investment management firms to:

> *to use commissions to purchase both execution and research services and charge this back to the client... (t)he execution fee would remain with the executing broker, while the non-execution fee would be placed in a CSA – an account from which the asset manager could pay any type of research producer, not just brokers*
>
> (Frost Consulting, 2012, p.7).

The CSA was designed as a device to promote competition and improve efficiency by separating research and execution payments. The introduction of CSAs has allowed fund managers to obtain a higher degree of independence in the advice they purchase, a wider choice of inputs and better value for money (Haig and Rees, 2015).

Despite these benefits, CSA adoption was slow. Even in 2013 it was estimated that CSAs accounted for less than half of the total market for research (Haig and Rees, 2016). Many firms maintained hybrid approaches, with CSAs complementing a conventional broker vote. Some asset management firms were, however, quick to try the new system. Evidence from agency brokers confirms that these early adopters were typically mid-sized UK firms who might not have been premium clients with top-tier brokerages. They may, as a result, have faced limited access to the more valuable parts of broker research.

Despite the requirements for careful monitoring of dealing commissions, FSA supervisory activity in 2011 found unacceptable compliance at 13 of 15 firms investigated. The FSA then reminded fund managers of the regulations and, after the FCA emerged from the FSA as a separate entity in 2013, entered a consultation process and review to take into account the views of stakeholders. In May 2014 the FCA consultation concluded with a clarification that client commissions should only be used for substantive research. The term clarification understates the importance of this communiqué: *'The FCA says it* *is a clarification. It is a clarification in the same way a dog shows its teeth when it does not want you to enter a house.'* (IMA Director of Legal and Compliance quoted in the Financial Times, 11 May 2014). In addition to the clarification, the CEOs of the largest 200 UK asset managers were required to personally attest that commissions have been spent with at least as much care as spending the firm's own money. This attestation has global ramifications given that many of the CEOs are based outside the UK and is therefore encouraged global adoption (CFA event, June 2014). We can see that the institutional changes in the UK aimed for value for money for the end investor. This contrasts with the US regulations which sought to curb misbehaviour. Even so, the changes envisaged in the UK will surely change analysts' incentives and may well be expected to subsequently affect behaviour.

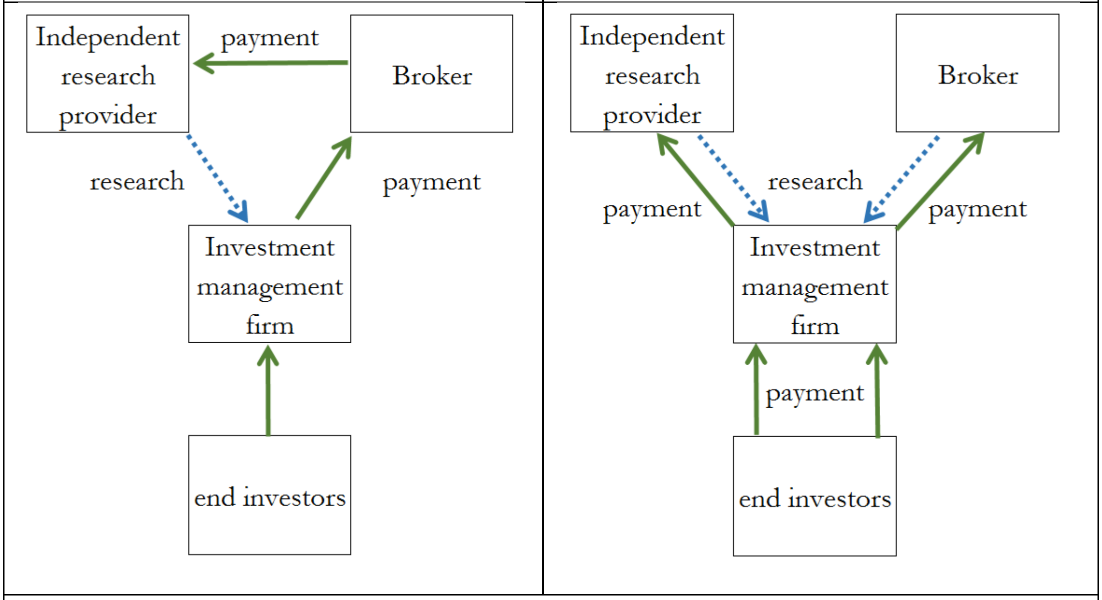
By 2016 it had become the norm (Unbundling Uncovered Event, 2016). Despite the likelihood that MiFID II would either prohibit or at best modify the CSA rules, some investment management firms strived towards 100% CSA adoption; several executives noted that adoption allowed them to document prices for research which had been less explicit in a traditional, percentage-based broker vote. The adoption of an invoice-based system created price records which could be presented to regulators or end investors to demonstrate orderliness. The invoices also provided management information. Analysis of ex-post prices could be useful in preparing to negotiate ex-ante prices. While many firms grappled with unbundling, one (buy-side) chief investment officer claimed that itemizing prices for research was a simple task because his firm had been an early adopter of CSAs (CFA Event, October 2016).

2.5.4 Mechanism 4: Research payment account (UK 2016 to present)

In 2016, the Financial Conduct Authority (FCA) introduced the concept of the Research Payment Account (RPA) (FCA, 2016). Under this system, buy-side firms can continue to use dealing commissions to pay for research, but an ex-ante research budget must be agreed with clients. The research budget and the account (RPA) which is used to disburse budget payments are independent of trading and therefore break the link between execution and research payments as shown in Figure 2d The RPA system requires the investment manager not only to be invoiced for research, as was the case with CSAs, but also for the end investor to be informed about the expected cost of research.

Figure 2d: Research Payment Account (RPA) Mechanism

Panel A: CSA	**Panel B: RPA**



Notes

The flow of information (research) is shown in dotted lines. Payment flows are in solid lines.

Panel A (left) recaps the method seen in the previous section.

Panel B (right) shows the Research Payment Account (RPA) where the investment manager negotiates a payment for research with brokers and research providers in the same way. This method was initially proposed by the FCA in 2015 and defined in subsequent statements in 2016 and 2017. Strict rules apply to the communication of RPA costs to end investors.

In practice, most UK investment managers and many other MiFIDII regulated firms have chosen to adopt the P&L system described in the following section.

Any investment management firm wishing to charge clients for research must disclose a budget to each client before receiving any research. The budget should reflect the volume, quality and cost of research required to meet the client's return objectives. In order for this budget to withstand the scrutiny of European regulators, it must be governed and reviewed by senior management. It must also be used to provide a regular summary of the research purchased and follow a clear audit trail. To meet the budgeted payments, an RPA must be created and funded by an explicit charge made to the end investor, thus ensuring that the research payment is not linked to trading activity in any way. RPAs can be funded by an explicit research charge to the end investor or by a CSA. In the event that less research is purchased than anticipated in the budget, the end investor must be reimbursed, or, if agreed, the balance could be used against the research budget in a subsequent period. Finally, the process to arrange these payments must be fully documented. Use of research with no explicit payment constitutes an inducement to trade under MiFID II. The wording of the directive prohibits the offering of gifts as the counter-gifts may reduce end investors' welfare.

RPAs also present the sell-side with considerable strategic and operational challenges (Quinlan, 2017). Strategic considerations include the type of products and services to be delivered, the industries and stocks to cover, the means of segmenting clients based on the amount they pay, and the provision of research to clients who are, or are not, subject to MiFID II. Operational concerns include contracts, billing and compliance; brokers need to be able to switch off the supply of research where it has not been requested. Participants report that MiFID II created confusion (RSRCHXchange, 2017, Figure 2.1). Unlike the classic strawberry market, which was seen as an opportunity to bring orderliness to the exchange process (Garcia, 2007), the introduction of MiFID II was perceived to be a disorderly process. Calls for clarity from regulators, a theme in all practitioner events I attended, persist even at the time of writing.

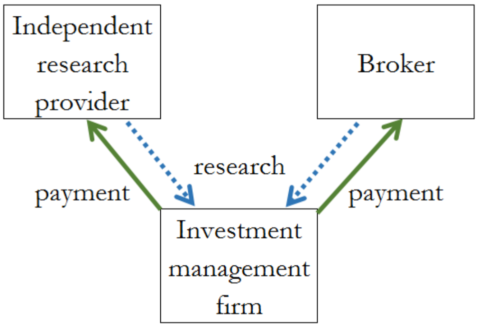
2.5.5 Mechanism 5: Profit and loss (P&L; UK 2016 to present)

Although MiFID II requires investment management firms to pay for research using a transaction-based system, RPAs are not the only option. Buy-side firms can instead bear the cost of research alongside other outgoings such as salaries, travel and data. The latter approach, recommended to the FCA in the early 2000s (Myners, 2001, Figure 2.1), has been most uncommon throughout the long history of analyst research.

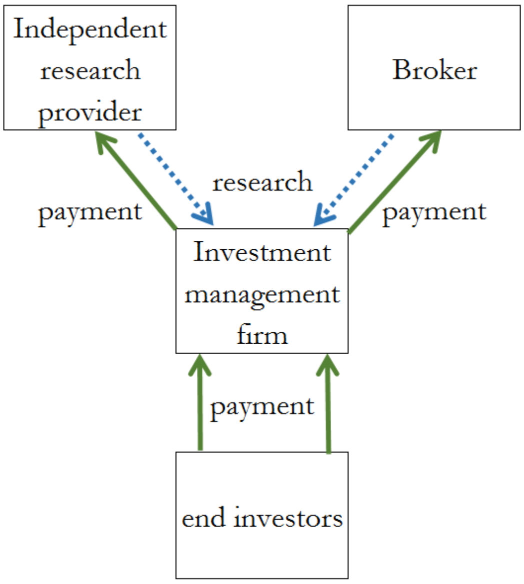
During 2017 the majority of firms decided to pay for research using the P&L method. While this will reduce a firm's profit, it was seen to reduce regulatory risk, compliance costs and internal distraction.

Figure 2e: Profit and Loss Payment Mechanism

Panel A: P&L



Panel B: RPA



Notes

The flow of information (research) is shown in dotted lines. Payment flows are in solid lines.

Panel A (left) shows the simplest method where investment management firms pay for their own research. The end investor makes no additional payment, although some investment management firms have sought to increase the fees they charge, i.e., incorporating the cost of research into the overall fee. This approach was uncommon prior to 2017.

Panel B (right) shows the Research Payment Account (RPA) where the investment manager negotiates a payment for research with brokers and research providers in the same way. This method was initially proposed by the FCA in 2015 and defined in subsequent statements in 2016 and 2017. Strict rules apply to the communication of RPA costs to end investors.

In practice, most UK investment managers and many other MiFIDII regulated firms have chosen to adopt the P&L system depicted in Panel A.

Myners (2001) called upon the UK investment management industry to bear the costs of research, and to reflect these in fees, rather than passing on the costs to end investors. Such a system would encourage the buy-side to use research efficiently and prompt the sell-side to offer a wider set of research services. This outcome appears to have been achieved indirectly, and rather abruptly, if belatedly, in the second half of 2017.

The neoclassical economic market model for investment research introduced by the UK regulator, as part of MiFID II, had a performative effect on the payment for research in the investment management industry. Since the effect has been to make the marketplace more like the economic model envisaged, we can say that this instance of performativity is of a Barnesian nature (MacKenzie, 2008).

It seems unlikely, however, that the investment management industry will entirely abandon its deeply engrained reciprocal system. The non-standardized, trust-based and relational nature of research suits a gift exchange system (Offer, 1997). Thus, it is likely that aspects of gift exchange will be retained. Buy-side firms with no client business in Europe need not comply with MiFID II and may choose to retain broker vote systems. This might include many hedge funds. An additional reason to expect less than universal adoption of new systems is simply that compliance by UK fund management and brokerage firms in this area seems to have been poor (CFA, 2014). The FCA has noted improvement in practice (Unbundling Uncovered, 2016). Over time, however, it seems likely that influential end investors, such as pension funds, will expect standardized approaches around the world regardless of MiFID II.

2.6 Conclusion

Using data collected from specialist industry events and documents, this paper identifies five market mechanisms used to pay for investment research in the past four decades. The characteristics of each mechanism are classified according to their orientation towards neoclassical market or gift exchange systems. The first two mechanisms feature no contracts or invoices. Here the price of research is not just opaque but often unknown both to supplier and consumer; the reciprocal mechanism exhibits characteristics of gift exchange. It is not, however, a free good. These systems prevailed throughout the present decade, often until 2017, and are still used in some firms outside the EU. A third system (CSA), which is invoice based, was made available in many markets between 2006 and 2017. Investment managers could elect to use this system to unbundle research payments from execution commissions. The third system revealed prices and removed the need to maintain a trading relationship with every research provider and afforded a wider selection of research, access to analysis from independent, non-brokerage firms and better value for money. Despite these benefits, the adoption of such methods was slow.

The fourth and fifth systems, almost unheard of until 2017, both require prices to be agreed and largely replace gift exchange systems. The key difference between these two current methods is whether the cost is charged to the end investor or the investment management firm. In 2016, buy-side firms were presented with the choice between some combination of RPA and P&L payment methods. Both systems require explicit prices to be negotiated and facilitate a market styled on neoclassical economic models. The majority of UK fund managers chose to change their approach and now pay for analyst research in the same way that most businesses would pay external consultants. Motivated more by uncertainty than by cost, investment managers have chosen to pay from their firms' own pockets; in doing so, they reduce, but do not eliminate, regulatory risks associated with the introduction of MiFID II. The end result is very similar to what might have been achieved by simply banning the use of commissions, as was proposed by the UK regulator in the early 2000s.

This shift in practice contrasts with a deeply engrained gift exchange system where investment managers recognize research in a similar way to the receipt of a gift and reciprocate with a counter-gift in the form of commission allocation. The counter-gift is proportional to trade size and therefore grows proportionally with assets under management. The relaxation of commission rules in the US and UK in the 1980s prompted no change in the marketplace design. In the immediate aftermath of the late 1990s dot.com bubble, the UK became known for regulatory attention to research marketplace efficiency. Even so, there was little change in practice. Broker votes persisted beyond the FCA's 2014 decision to support more demanding rules on research payment as part of MiFID II. By 2016, it had become clear that the gift exchange system would no longer be compliant after the introduction of the new directive from the start of 2018.

The change has important ramifications. Investment managers must now evaluate the cost of research alongside their other costs incurred to run the business, allowing managers to directly compare external and internal research costs such as analyst compensation. Greater scrutiny can be expected than was typical in many broker vote systems. Research which is not perceived to add value will be removed from future research budgets. In many cases the overall budget for research will be lower than it was prior to 2018. Research providers must negotiate the price of research in monetary terms rather than percentage points and in advance of use. For many firms this will be a significant change. Providers will have a clearer view of their demand schedule. Few brokers now attempt to provide full coverage of listed stocks in all sectors or countries; instead, most specialize. End investors have greater transparency regarding the costs of investing. Since fund managers have a regulatory requirement to avoid paying for underutilized research, we can expect less duplication. Competition should result in greater value for money. It may be harder for investors to find and evaluate research but entrepreneurs have set up stall to introduce providers, curate research and evaluate research efficiency

Table 2.1: Schedule of events attended

Event title	Organization	Date	Notes
The Extinction of Sell-Side Research	CFA UK	May 2013	Presentation by Frost Consulting on the economics of research based on data collected through consultancy
Research Review Panel Discussion	IMA	July 2013	Launch of IMA report in response to FCA thematic review on paying for research
The Future of Equity Research	CFA UK	September 2013	Panel discussions on challenges facing the existing model of research, changing industry structure (chaired by author) and research content
Professionalism Conference	CFA UK	April 2013	Presentations by fund management and FCA leaders on conflicts of interest, professionalism and firm culture
The Changing Market for Investment Research	CFA UK	October 2014	Presentation by author to CFA Society of the UK (Research Analysts Special Interest Group)
Broker Vote Solutions	Bloomberg	December 2014	Event for Bloomberg buy-side client. Author's survey conducted.
Unbundling Uncovered	Substantive Research	November 2015	Full-day conference attended by over 200 buy-side, sell-side and independent research specialists

Paying for Research	CISI	February 2015	Presentation by author followed by discussion with 20 CISI Fellows
Research Marketplace	CFA UK	October 2016	100 investment professionals, staff and students
Research Marketplace	Scottish Competition Forum	October 2016	20 law firm partners, competition economists and University of Edinburgh School of Law academics
Fixed Income Research	Unbundling Uncovered	October 2016	280 fund management/investment bank executives
Paying for Research	Australian Securities and Investments Commission	November 2016	(Sydney) 30 staff (plus 12 more via video link to Melbourne, Perth, Adelaide, Hobart and Canberra)
Unbundling Uncovered II	Substantive Research	November 2016	Full-day conference attended by over 250 buy-side, sell-side and independent research specialists
Investment Research	Institutional Investor Fixed Income Forum	March 2017	(Madrid) Over 100 fixed income and multi-asset fund managers
MiFID II Research Implementation	Institutional Investor Nordic Forum	March 2017	(Stockholm) Over 100 pension fund executives (asset owners) and fund managers
Unbundling Uncovered II	Substantive Research	November 2018	Full-day conference attended by over 330 buy-side, sell-side and independent research specialists

This table reports conferences and debates attended by the author. In most instances, these were organized by professional bodies representing market practitioners and in one instance, the Bloomberg meeting, the event was organized by a data provider together with input from the author.

Figure 2.1: Industry and professional documents

CFA UK (2013) The Future of Equity Research (event summary made available in Professional Investor) October 2013. Author's copy. Available to CFA UK members.

CFA UK (2014) Response to FCA Consultation CP13-17: Consultation on the use of dealing commission rules.
http://www.frostconsulting.co.uk/futureofresearch.php

CSFI (2017) A Level Playing Field for Investment Research? Challenges facing the buy-side, sell-side and independents
http://www.csfi.org/20171024-a-level-playing-field-for-investment-research

Euro IRP (2018) Members Study
https://euroirp.com/wp-content/uploads/2018/08/Euro-IRP-FINAL-Members-Study-Report-2018.pdf

Extel (2011) Buy-side Broker Reviews: Commission or Omission
http://www.extelsurvey.com/ (last accessed 15/12/2013)
FCA (2013) Consultation Paper 13/17: Consultation on the use of dealing commission rules
http://fca.org.uk/your-fca/documents/consultation-papers/cp13-17

FCA (2014) Discussion Paper 14/3: Discussion on the use of dealing commission regime
http://www.fca.org.uk/static/documents/discussion-papers/dp14-03.pdf

FCA (2016) Consultation Paper 16/29: Markets in Financial Instruments Directive II Implementation – Consultation Paper III
https://www.fca.org.uk/sites/default/files/cp16-29.pdf

Frost Consulting (2012) Commission Sharing Arrangements: The changing face of research procurement
http://www.frostconsulting.co.uk/publications.php

Frost Consulting (2012) Introduction to Commission Unbundling
http://www.frostconsulting.co.uk/publications.php

Global Analysts Research Settlement (2003)
http://www.sec.gov/news/speech/factsheet.htm

Investment Association (then known as IMA) (2013) The Use of Dealing Commission to Purchase Investment Research
https://www.theia.org/industry-policy

Investment Association (2014) Response to FCA Consultation CP13-17: on use of dealing commission rules
https://www.theia.org/industry-policy

Kay, J. (2012) The Kay Review of UK Equity Markets and Long-Term Decision Making
https://www.gov.uk/government/news/kay-review-publishes-report-on-uk-financial-sector

Myners, P. (2001) Institutional Investment in the United Kingdom: A Review Report to HM Treasury.
http://webarchive.nationalarchives.gov.uk/+/http://www.hm-reasury.gov.uk/media/1/6/31.pdf

Oxera (2009) The impact of the new regime for use of dealing commission: post-implementation review
http://www.oxera.com/Oxera/media/

Quinlan (2016) Research in an unbundled world
https://www.quinlanandassociates.com/insights/

RSRCHXchange (2017) MiFIDII Survey Summary Report
https://www.rsrchxchange.com/regulation/

All reports last accessed 26 June 2019

Figure 2.3: Timeline for each payment method

	On execution	Broker vote	CSA	RPA	P&L
1920–1949	Yes				Yes*
1950–1975	Yes				Yes*
1975–1999	Yes				Yes*
2000–2005	Yes	Yes			Yes*
2006–2015	Yes	Yes	Yes		Yes*
2016–2017	Yes	Yes	Yes	Yes	Yes*
2018 onwards (MiFID firms)				Yes	Yes

The typical use of payment methods in the UK and US stock markets since the brokerage industry and the financial analyst emerged in the 1920s are shown in five columns.

All methods could be adopted during calendar years 2016 and 2017. The RPA was not fully defined until later in 2016. Consequently, there was a short window in which buy-side firms could decide which approach to adopt. Recently, a short period existed in which all methods were allowed: 2016–2017. In 2017 the majority of UK fund managers announced adoption of P&L for equity funds.

*P&L was available to investment management firms prior to 2018 but was uncommon.

CHAPTER 3

Mind versus machine: analyst and machine learning equity valuations

3.1 INTRODUCTION

In this paper, we investigate the ability of machine learning (ML) to mimic analysts' equity valuations. Specifically, we compare the information content of analyst and ML valuations. The adoption of ML algorithms which can automatically re-estimate relationships in large, high dimensional datasets have become a major theme in investment management. For example, a search of the Financial Times website reveals over 1,000 articles relating to ML. Given the profound impact this technology is expected to have on the industry, this level of attention is understandable. Many articles advocate awareness, or adoption, of ML in investment management but few evaluate its effectiveness. Our paper takes a step towards spanning this gap in the literature, specifically by examining the potential for ML to perform analyst research.

Several important debates motivate our study. Although analysts have been shown to have an important role in supporting the informational environment (Brown et al., 2015), their ability to make accurate predictions remains questionable (Bradshaw et al., 2017). Sell-side analysts, for many decades the most typical and available analyst category, face adverse incentives, which result in optimism bias which interferes with accuracy. All brokerage analysts, whether affiliated to an investment bank or not, can be expected to face a trade-off between maximising reputation, which might be demonstrated by predictive accuracy, and encouraging institutional investors to trade, e.g., through overoptimistic buy recommendations. Recent regulation has required unprecedented change in the supply of investment research for investment firms operating in 31 European Economic Area countries. Since January 2018, MiFIDII has mandated that research must be invoiced separately from dealing and that investment managers cannot use research which they have not explicitly paid for. Brokerage firms must now sell research in a similar way to independent research providers (IRPs), e.g. through quarterly subscriptions. Prior empirical evidence on analyst bias and incentive has become less relevant.

Firms providing analyst research are responding to MiFIDII amidst a shift towards passive investing, which requires little or no research, and falling commission rates, which reduce brokerage revenues. The demand and revenue for research has therefore been under pressure due to structural change in the last decade. While media reports of “bloodbath” and “exodus” (Gordon, 2017) in sell-side roles may contain a modicum of hype the scale and stature of sell-side research has diminished (Guan et al., 2013; Groysberg and Healy, 2013; Haig and Rees, 2016). The use of ML is becoming more common in knowledge industries and it would not be surprising to find analysts in some firms making more use of ML, and in other firms there may even be attempts to replace humans with automated alpha-generating algorithms.

Our setting provides a rare opportunity to compare the forecasting ability of analysts versus ML. Such opportunities are unusual because investment firms, whether buy-side, sell-side or independent, seek to use the technology for their private advantage. Our sample comprises some 100 analysts at a well-established firm which supplies investment managers with analyst outputs on a subscription fee basis. Since the research they produce is used only by external investors, and is paid for by fees rather than brokerage commissions, they are free from important adverse incentives which can impede the objectivity of their sell-side counterparts. This setting is unlike most prior studies, which use brokerage analysts, yet is similar to the post-MiFIDII landscape for analysts. It therefore gives us a timely preview of research in the years to come.

Regressions on valuations of US and international firms show that analyst valuations contain information which is not fully encapsulated by valuation multiples, size, risk, momentum, profitability and the I/B/E/S consensus target price but with minimal explanatory power when used alone. ML does not mimic analysts but instead acts as a significant contrarian signal. Investors could use a combined signal augmented with price momentum but they are unlikely to invest on the basis of a contrarian signal without a compelling explanation. More importantly, our analysis points to how ML might be adapted into a reliable non-contrarian indicator, for example by including a momentum factor, or by reducing the sensitivity of ML valuation to price changes.

Our primary contribution is evidence on the efficacy of ML in investment analysis. This evidence is timely given the current focus on financial technology known as FinTech. Additionally, we contribute to the narrow strand of literature on independent investment research. This is especially important because, as chapter 2 explains, MiFIDII requires most brokerage firms to charge for research separately from dealing commissions. Since our sample analysts make unbiased predictions, we have a near-ideal setting in which to examine independent analysts' valuations, and this provides a preview of what to expect since MiFIDII took effect in 2018. Finally, since our sample is global, yet contains a sizeable US sample, we provide new evidence of both ML adoption and independent research in global markets, yet still allows us to make comparisons with existing research which tends to be set in the US. We expect to see more research on these themes.

3.2 BACKGROUND

3.2.1 THE USE OF MACHINE LEARNING IN SYSTEMATIC INVESTMENT MANAGEMENT

Active investment management firms come in a variety of forms. Equity investment funds have historically been classified as fundamental or quantitative. These labels can easily be contested as quantitative funds often use fundamental data as inputs (and are sometimes labelled "quantamental"); equally, many fundamental managers are quantitatively sophisticated.

Fundamental managers may also object to being labelled as “traditional” and see their approach as innovative. Lopez de Prado (2018) classifies strategies into discretionary, where investment managers make decisions based on judgement and experience, and systematic approaches, where decisions are algorithmic. We follow this distinction.

3.2.2 MACHINE LEARNING

ML can be used in discretionary and systematic investment strategies and take on different meanings depending on the domain of use. Empirical analysis in finance has tended to use econometric approaches which predate ML and require the analyst to estimate a model to measure the relationship between variables. Investors can update this model when new data arrives or at periodic intervals in order to make investment decisions (systematic investment managers), and academics often simulate such strategies, for example by testing the performance of portfolios representing company characteristics. ML can also support such decisions (discretionary investment managers), for example by processing unstructured text and image data to provide business analysis to investment managers. ML differs in that the algorithm, rather than the analyst, determines the model. Data scientists tend to refer to this as “training” rather than estimation. For more detail, and specifically for readers with a background including economics, Athey and Imbens (2019) provide a comparison between machine learning and established econometric methods.

Gu et al. (2019) define machine learning in the asset pricing domain to capture approaches which firstly, offer the flexibility to analyse large datasets; secondly, guard against overfitting and thus reaching stable out-of-sample prediction; and thirdly, where there is a practical method to reach the best model among multiple near-optimal solutions where the optimal model is unknown or may be impractical to reach. On this last point, the issue is that it may not be feasible to compute the exact optimum due to cost or time required to evaluate every possible model generated by the algorithm. ML can be classified as supervised, where a set of variables or features are provided (such as random forest and neural networks), and unsupervised ML where the algorithm determines relationships between observations (e.g. cluster analysis). The ML in our setting takes the former approach, i.e., it is supervised.

The type of ML chosen by the firm in out setting is called random forest, which is built on the concept of decision trees. A decision tree plots a sequence of steps. Decision trees can be used for classification or prediction; in our setting the latter is the goal. Data scientists call such trees regression trees. This can be confusing in finance as the approach is quite different from regression analysis in standard econometrics.

In the first step of a decision tree, observations are partitioned into two bins (also called nodes or leaves) using an objective function which minimises the probability of a random observation being classified in the wrong bin. The partition is made using a threshold value of one of the explanatory variables (features). There are now two bins. (It is also possible to have more than two but two is simpler to start with two). The second step is to partition the data in each of the new bins, again using the objective function to each bin. These two new decisions are most likely based on threshold values of different variables. We now have a second level of the tree featuring four bins. We could continue to “grow” the tree down the page; the graphical representation can be likened to an upside-down tree with branches reaching downwards. At some stage we can decide to stop adding levels. The number of levels is referred to as the depth of the tree. Suppose we stop with the two levels described above, i.e., at a depth of two. This means we have four buckets. The average value of the MVP in each bucket would be the predicted value for any stock falling into that bucket, i.e., with characteristics which meet the threshold requirements to end up in that bucket. The decision tree now serves as a model which we could apply to firms which were not included in the original sample.

We could stop here and use the outcome of the decision tree as the forecast. One sell-side firm supplied such a model to investment managers in the 1990s. To improve the prediction accuracy Breiman (2001) proposes the use of a large number of trees, i.e., an ensemble, which is typically known as a random forest. Ensemble approaches are helpful when there is a low signal to noise ratio resulting in marginal predictive power for any single tree. Using a bootstrapping approach, trees can be grown from random subsamples of the dataset. The firm in our setting generates 500 decision trees and the ML market valuation to price is the average of these valuations for any gives share. Random forest has become something of a standard method in finance. Informal conversations with practitioners working in this area indicate that the method is suitable because it is relatively simple and transparent. Gu et al. (2019) include it in their comparisons and find that ML simpler methods with few levels (shallow) tend to outperform complicated with many levels (deep), such as deep neural networks.

3.2.3 SYSTEMATIC EQUITY STRATEGIES

Systematic use of reported accounting information has been shown to outperform analyst consensus forecasts. Financial statement analysis of value and growth firms can be used to devise compelling investment strategies (such as Ou and Penman (1989), Frankel and Lee, (1998) Piotroski (2000) and Mohanram (2005). Wahlen and Wieland (2011) use simple financial ratios to predict earnings with greater accuracy than the sell-side analysts' consensus. Importantly, the models do not rely on forecasts. To provide some concrete examples, Piotroski (2000) predicts the returns of value (high book to market) stocks using the variables such as cash flow from operations, debt to total assets, gross margin, asset turnover and return on assets; this strategy was adopted, for example, by the brokerage arm of Societe Generale. Wahlen and Wieland (2011) build a simple model using a similar selection of variables to predict company earnings with greater accuracy than the sell-side analysts' consensus. Importantly, the model does not rely on forecasts. As such, these approaches present a systematic alternative to analyst-based approaches.

Systematic equity strategies have typically relied upon factors such as book to price, past returns, firm size and earnings quality; the models apply and extend well-cited asset pricing models such as Fama and French (1992) and Carhart (1997).

Quantitative analysts estimate models on as many prior periods as possible in the design and testing phases. Once live, the data inputs to the model are updated regularly, usually at daily or monthly intervals. Estimation of the model is typically infrequent due to the scale of effort required. Quantitative analysts might re-estimate the model from time to time, for example to incorporate a new variable or simply because investment performance has been unsatisfactory. Systematic investing has tended to use models which are static in the sense that there is no systematic schedule for re-estimation. Examples of firms established in such approaches are AQR, LSV and the Barclays Global Investors business acquired by Blackrock in the late 2000s (Fabozzi et al., 2010).

In the past decade, systematic investing has shifted from static processes to dynamic models. In a survey of investment managers (Fabozzi et al., 2010), respondents stressed the need for dynamic strategies in preference to static models. More recent approaches seek to use ML to forecast earnings (Ball and Ghysels, 2017), target prices and valuations (Morningstar, 2013) or to forecast stock prices directly (Fabozzi et al., 2010). Even if ML seems not to have become dominant in mainstream quantitative investing (Kahn, 2018), a recent study shows signs of significant innovation in investment research (Grennan and Michaely, 2018). While such innovation might be seen as a contest between “mind” and “machine”, one recent study (Ball and Ghysels, 2017) moves beyond direct comparison and demonstrates that a blend of ML and analyst approaches is optimal.

Our search for prior studies of ML in investment research therefore reveals that while practitioners invest heavily to forge ahead with ML applications, there is an important gap in the literature. Fortunately our dataset provides a near-ideal setting. Our data comprises valuations made by over 100 analysts at a well-established IRP. These valuations are expressed relative to the current share price. We label the analyst valuation to price ratio *AVP.* The same firm employed ML to mimic *AVP* and we label this ML valuation to price ratio *MVP.* Both valuations were made available to Morningstar’s global client base comprising thousands of buy-side firms and therefore provide an opportunity to evaluate an early adopter of ML-based investment research.

Analysts have been shown to play an important role as information intermediaries for many years (Beaver, 2002; Bradshaw et al., 2017).

3.2.4 Target prices and discretionary investing

Analysts' target price forecasts have been shown to be more informative than earnings estimates or recommendations (Brav and Lehavy, 2003). Even so, the literature has developed to show that they are "optimistic, inaccurate, and of little long-run investment value" (Joos et al., 2016, p. 645). Analyst target price accuracy is greater where valuation approaches such as discounted cash flow (DCF) models are chosen (Demirakos et al., 2010; Gleason et al., 2013).

Sell-side analysts seek to influence discretionary investment managers in order to attract brokerage commissions. Prior studies have examined the accuracy of analysts' predictions; the large literature on earnings forecasts and stock recommendations has been complimented by other key ouputs such as target prices. Target price can be the analyst's expectation of a given firm's stock price, i.e. valuation, in 12 months' time. In this chapter, we examine valuations made by analysts for each stock in their coverage list. In short, while there is evidence that analyst target prices have some information, in practice it is difficult to extract this using systematic investment strategies.

Several studies have compared independent research provider (IRP) analysts to sell-side analysts. Many of these studies compare the bias or accuracy of independent analysts who are not subject to conflicts between the best interests of their investment management clients and the investment banking department (Mehran and Stulz, 2007). Such conflicts were common the Global Settlement which, in 2003, restricted the links between analysts and investment banking activities (Bradshaw et al., 2017). A long line of research stems from this: for example, Jacob et al. (2008) and Ertimur et al. (2007) are among those to find earnings forecast accuracy to be positively related to the profit made by the brokerage if the analyst is not incentivised by investment banking. These studies were timely in the aftermath of the 2003 Global Analysts Research Settlement.

There is, however, a more fundamental conflict which the literature tends not to address. Brokerage firms generate their revenue in the form of commissions and so their analysts remain incentivized to induce trading and this stands regardless of whether investment bank affiliations exist. Analysts can only be expected to be objective if they do not face incentives to encourage investment managers to change their view and therefore make trades.

Prior studies have often muddied waters by including unaffiliated brokers who still face important conflicts: Barber et al. (2007) are forced to blend unaffiliated analysts with IRPs due to the paucity of truly independent firms. MiFID II has increased attention on inducements to trade. This makes our study timely.

IRP studies have tended to focus on recommendations and as a result we know little about IRP target prices. Sell-side target prices have been shown to convey incremental information, i.e. not contained in earnings forecasts or recommendations, for medium-term investment horizons (Brav and Lehavy, 2003) even if substantial optimism bias persists (Bradshaw et al., 2013) long after the Global Settlement. Allee et al. (2017) document that US analyst recommendations made by a large IRP could have been used to generate positive alpha. Our study extends this work by providing international evidence and comparing analyst valuations with ML counterparts. Our work differs in that we focus on analysts' valuations, a more granular output which has potential to convey more detailed information such as level of conviction.

Analyst predictions and recommendations may be most valuable for small cap firms, which have been shown to be more responsive to new recommendations (Womack, 1996); most likely because of the public information available to small cap investors is less complete (Stickel, 1995). Analysts can therefore play a larger role in the dissemination, especially in the case of analysts at larger brokers. Barber et al. (2001) compare portfolios of stocks with buy and sell recommendations; the abnormal returns generated on small-cap portfolios is largest which adds further evidence to the enhanced role of analyst information on smaller companies.

The type of ML chosen by the firm in our setting is called random forest, which is built on the concept of decision trees. A decision tree plots a sequence of steps. Decision trees can be used for classification or prediction; in our setting the latter is the goal. Data scientists call such trees regression trees. This can be confusing in finance as the approach is quite different from regression analysis in standard econometrics.

In the first step of a decision tree, observations are partitioned into two bins (also called nodes or leaves) using an objective function which minimises the probability of a random observation being classified incorrectly The partition is made using a threshold value of one of the explanatory variables (features). There are now two bins. (It is also possible to have more than two, but it is simpler to start with two.) The second step is to partition the data in each of the new bins, again using the objective function. These two new decisions are most likely based on threshold values of different variables. We now have a second level of the tree, featuring four bins. We could continue to 'grow' the tree down the page; the graphical representation can be likened to an upside-down tree, with branches reaching downwards. At some stage we can decide to stop adding levels. The number of levels is referred to as the depth of the tree. Suppose we stop with the two levels described above, i.e., at a depth of two. This means we have four buckets. The average ML valuation in each bucket would be the predicted value for any stock falling into that bucket, i.e., with characteristics which meet the threshold requirements to end up in that bucket. The decision tree now serves as a model which we could apply to firms that were not included in the original sample.

We could stop here and use the outcome of the decision tree as the forecast. One sell-side firm supplied such a model to investment managers in the 1990s. To improve prediction accuracy, Breiman (2001) proposes the use of a large number of trees, i.e., an ensemble, which is typically known as a random forest. Ensemble approaches are helpful when there is a low signal to noise ratio, resulting in marginal predictive power for any single tree. Using a bootstrapping approach, trees can be grown from random subsamples of the dataset. The firm in our setting generates 500 decision trees, and the ML market valuation to price is the average of these valuations for any given share. Random forest has become something of a standard method in finance. Informal conversations with practitioners working in this area indicate that the method is suitable because it is relatively simple and transparent. Gu et al. (2019) include it in their comparisons and find that simpler ML methods, with few levels (shallow), tend to outperform complicated methods with many levels (deep), such as deep neural networks.

ML has been of interest to researchers for many years. Early studies proposing the use of neural networks (see, for example, Swales and Yoon, 1992; Kryzanowski et al. 1993) show that this topic was of interest to investment professionals even in the early days of vendor data on analyst forecasts. Bouwman et al (1987) is an example of multidisciplinary research on expert systems, where the authors seek to design an expert system based on qualitative data from structured interviews with financial analysts. To our knowledge there was little progress in research in these areas, a situation reflected in practice. Surveys of investment approaches such as Fabozzi et al. (2010) and, more recently, Kahn (2018) reveal few instances of ML and none which obviously seek to replicate analysts outputs or to extend analyst coverage. This situation seems to have changed in the last few years, when many investment management and research firms published white papers, albeit with sparse if any empirical evidence to support the merits of the new technology adoption. The proprietary nature of the data has prevented scholarly research except where researchers attempt to develop their own ML methods to predict earnings (see for example Ball and Ghysels, 2017) or stock returns (Gu et al. 2019).

3.3 HYPOTHESES

Our study compares the informativeness of analyst and ML equity valuations. We do this using valuation estimates as target prices, thought to be the most valuable summary of an analyst's research on a given company (Dechow and You, 2017). Although target prices contain information which is not impounded in earnings forecasts or recommendations (Brav and Lehavy, 2003), it is difficult for investors to create systematic investment strategies based on these alone (Bradshaw et al., 2017). This is not surprising as the efficient market hypothesis tells us that we cannot expect analysts to forecast share prices with ease. There is, however, some evidence that independent, unbiased or experienced analysts may have some advantages in the search for alpha. Barber et al. (2007) confirm that the buy recommendations of affiliated brokers are substantially less informative than those of unaffiliated or independent analysts. Independent analysts may, however, have less experience and more limited access to resources and may in fact have similar optimism bias to sell-side analysts (Bradshaw et al., 2017).

It is, therefore, important for us to consider the experience, bias and valuation approach of our sample analysts. Taking experience first, our sample analysts are experienced and work in an established franchise. Second, and perhaps more importantly, they produce a near-symmetrical distribution of recommendations which is evidence of minimal, if any, bias. Analysts with unbiased recommendation distributions have been found to be more accurate (Barber et al., 2006). Finally, their use a systematic valuation methodology (Brown et al., 2015) improves their chances of accuracy compared to analysts who rely on multiples (Gleason et al., 2013).

Around one third of the firms analysts are CFA charter holders, a designation which requiring training, experience and adherence to a code of ethics. Kang et al. (2018) find analysts who have earnt this designation to make more accurate predictions.

3.4.1 First hypothesis: incremental information in analysts' valuations

Bradshaw et al. (2013) establish that target prices are more informative than earnings forecasts or recommendations. This evidence is, however, made from sell-side analyst samples. Independent analysts may be more accurate, since they do not face incentives to generate trades. Alternatively, they have sometimes been found to have less experience and access to resources and so may be less accurate. Our key variable of interest is the forecast fair value to price ratio issued by analysts and the ML algorithm and functions as a target price (for example, it is presented as such to I/B/E/S). Morningstar analysts are free from the conflicts inherent in brokerage firms and so can focus on accuracy. Any bias in their predictions is most likely to be due to behavioural traits.

The distribution of recommendations and target prices issued by the firm are also remarkably symmetrical compared to the typical buy-oriented distribution. Barber et al. (2006) find that analysts at firms with more evenly distributed recommendations tend to be more accurate. Allee et al. (2017) document the positive predictive ability of Morningstar analysts. Our first hypothesis is:

> ***H1: Independent analysts' valuations contain incremental information about subsequent stock returns***

3.4.2 Second hypothesis: incremental information in ML valuations

The extant literature provides some evidence that accounting variables can produce abnormal returns (see, for example, Piotroski, 2000) which are superior to returns based on analysts' consensus recommendations (Wahlen and Wieland, 2011). ML developed by a well-regarded supplier might be able to improve upon such simple models. We hypothesize that ML will show some ability to predict stock returns. Our second hypothesis is:

> ***H2: ML valuations contain incremental information about subsequent stock returns***

Given the range of prior evidence on target price accuracy, it is also reasonable for us to expect differential predictive power for both sources of forecasts. Prior evidence on ML and analysts' earnings predictions indicates that there may be complementarity between the two (Ball and Ghysels, 2017). A combination of analyst and ML valuations might be most effective in our setting too.

3.5 RESEARCH SETTING, DESIGN AND DATA

3.5.1 Morningstar analyst outputs

Morningstar analysts follow a structured approach to fundamental analysis. Morningstar requires analysts to score each company on a list of criteria: competitive positioning, financial health, stewardship and fundamental risk. The cash flow projections and the assumptions emerging from this process are plugged into a proprietary DCF model which is adjusted for industry- and company-level assumptions; such an approach is associated with superior analyst forecasts (Gleason et al., 2013). Analysts have some discretion to triangulate the DCF valuations using multiples but are required to use the DCF. Even so, the process is more standardized than many brokerage firms, both geographically and across industries, and places an emphasis on detailed valuations rather than the more common approach of selecting multiples (Brown et al., 2015; Pinto et al., 2015).

Our sample has two unusual qualities. First, analysts are free from many of the incentive misalignments which are typical in brokerage firms (unlike the sample of Barber et al., 2007). We can therefore expect any bias or systematic error to be down to traits of human behaviour. Second, they generate a balanced distribution of valuations and recommendations, with approximately the same number of buy and sell recommendations, and an even greater proportion of neutral views. Any bias should therefore be due to human behaviour rather than incentives. This makes our sample ideal for studying analyst behaviour.

3.5.2 Morningstar ML outputs

Morningstar's ML process is explicitly designed to expand coverage. By 2017 the firm claimed to value over 50,000 stocks using the algorithm. The quantitative fair value to price estimate, the primary output, is designed to be analogous to the analyst fair value estimate. Rather than replicate the processes used by analysts, the algorithm is designed to forecast valuation using the characteristics of overvalued or undervalued firms. By measuring the impact of these characteristics on analysts' valuations, the algorithm attempts to value stocks regardless of analyst coverage. ML inputs are based on financial statements and market data, not forecasts made by analysts. The model is updated and re-estimated on a daily basis. More detail can be found in Morningstar (2013).

The ML algorithm is supervised using a predetermined input set comprising a typical set of accounting variables such as (return on assets, earnings yield, sales yield, book yield and revenue), market variables (volume, market capitalization and the largest peak-to-trough drawdown in each trailing 12-month period). The last is less common in equity quant models than a price momentum measure.

Table 3.1: Inputs to the ML algorithm

Variable	
Return on assets	Trailing 12 months
Earnings to price	Trailing 12 months
Sales to price	Trailing 12 months
Book value to price	Most recent
Equity (share price) volatility	Trailing 12 months
Maximum drawdown (maximum fall from peak to trough)	Trailing 12 months
Total revenue	Trailing 12 months
Market capitalization	Most recent
Enterprise value	Most recent
Average daily volume	Trailing 12 months
Enterprise value to market value	Trailing 12 months
Sector	Most recent

Source: Morningstar (2013). Data sources are not disclosed; we assume the data is sourced from Morningstar's own database. Our own specifications of these variables are shown in table 4.2.

The ML valuation is the mean of 500 estimates of the relationship between input variables and analyst valuation to price estimates. Each iteration uses a random subset of the available input observations. This process, which is known as a random forest, is designed to mimic analyst valuations. It updates both the data and the estimation automatically on a daily basis. It learns based on analysts' predictive ability.

3.6 METHOD

Prior research on target prices has tended to use a cross-section of sell-side firms rather than a single-firm study (see, for example, Bilinski et al., 2012). Some studies, particularly those which compare buy-side analysts with their brokerage counterparts, delve into a single-firm case study (the appendix to this chapter contains a table of related studies). Our setting allows us to make a direct comparison between the valuation to price forecasts of Morningstar's analysts and ML algorithm; for this reason, we also use a single-firm study. First we examine the determinants of each valuation measure by performing the regression shown in **Equations 1A and 1B**.

3.6.1 Determinants of analyst fair value to price

Morningstar discloses the accounting and market-based inputs to their ML algorithm (Table 3.1). Although we do not know the relative weighting given to each, which will change daily as the ML is re-estimated, we can expect *MVP* to be empirically determined by these factors. To understand determinants of *AVP* and *MVP* we use the following regressions:

Equation 1A:

$$AVP_{i,t} = \alpha + \beta_1 RETURN_{i,t-1} + \beta_2 IBES\_TPP + \beta_3 MVP_{i,t} + \beta_4 ROA_{i,t} + \beta_5 BTM_{i,t} + \beta_6 FIRMSIZE_{i,t} + \beta_7 VOLUME_{i,t} + \beta_8 LEVERAGE_{i,t} + \beta_9 BETA_{i,t} + \beta_{10} VOLATILITY_{i,t} + \varepsilon_{i,t}$$

Equation 1B:

$$MVP_{i,t} = \alpha + \beta_1 RETURN_{i,t-1} + \beta_2 IBES\_TPP + \beta_3 AVP_{i,t} + \beta_4 ROA_{i,t} + \beta_5 BTM_{i,t} + \beta_6 FIRMSIZE_{i,t} + \beta_7 VOLUME_{i,t} + \beta_8 LEVERAGE_{i,t} + \beta_9 BETA_{i,t} + \beta_{10} VOLATILITY_{i,t} + \varepsilon_{i,t}$$

3.6.2 Variables

We obtain the complete set of Morningstar analyst and ML valuation estimates, *AVP* and *MVP* respectively. After merging with Worldscope and Datastream variables, our dataset comprises some 17,000 firm/quarters of analyst and ML forecasts, both provided by Morningstar. This sample is somewhat larger than the single-firm studies listed in the Appendix.

Table 3.2: Variable definitions

Name	Definition	Source
AVP	Analyst valuation estimate divided by latest price	Morningstar
MVP	Machine learning (ML) algorithm valuation estimate divided by latest price	Morningstar
IBES_TPP	I/B/E/S consensus target price for each firm divided by current price at the end of each quarter	Datastream
RETURN	Total return calculated as the change in return index (RI)	Datastream
BETA	Bayesian adjusted beta = 0.4 + 0.6 * beta estimated from 5 prior years of monthly data. Estimates below zero or above 2.25 are set to missing. We use S&P1500 index for US firms, otherwise AC World ex-US index	Datastream
FIRMSIZE	Firm capitalization measured by market capitalization of equity in US dollars	Datastream
LEVERAGE	Leverage is measured as total debt to market capitalization of equity at quarter end. Estimates above 3 are set to missing	Worldscope
BTM	Book to market is book value of equity divided by market capitalization of equity at quarter end. Estimates below -0.5 or above 4 are set to missing	Worldscope
INDUSTRY	GICS level 2 industry classification	Datastream
COUNTRY	Country of listing	Morningstar
ROA	Worldscope total assets and Datastream market value	Datastream
VOLATILITY	Volatility is the standard deviation of 13 trailing weekly returns, including dividends, in local currency and annualized by multiplying by the square root of 52. Estimates below 0.1 or above 0.75 are set to missing	Datastream
SP	Revenue divided by market value	Datastream
BP	Book value of equity divided by market value	Datastream
REVENUE	Total sales	Datastream
EV	Book value of total debt plus market value	Datastream
EV/MV	EV divided by market value	Datastream
VOLUME	Average number of shares traded in the past quarter	Datastream

RETURN is the percentage change in Datastream total return index over each calendar quarter. Datastream is recognized as one of the most reliable providers of total returns data for international samples. To handle extreme values, we follow the approach of Ince and Porter (2006) and remove returns exceeding 300%. Our sample of firms is determined by Morningstar analyst coverage and all firms entering our sample are listed. This mitigates the likelihood of extreme returns, which are often associated with small companies. Return is the dependent variable in our regressions; we also include lagged return as momentum factor as is typical in the asset pricing literature and in practice.

Morningstar's ML is unusual as it does not contain momentum but does include a drawdown measure representing the greatest change from peak to trough based on the 12 month-end prices in each preceding quarter. We choose three-month realized returns as this is a more complete momentum measure and one commonly used by scholars and practitioners; replacing this with a 12-month version makes little difference to our results.

AVP is the valuation estimate made by the Morningstar analyst divided by the current price as at the end of each quarter.

MVP is the ML valuation to price estimate made by ML at the end of each quarter.

We also include a range of control variables, most of which are typical when examining the cross-section of stock returns. Reported financial statements are used to compute accounting variables. We use data from annual financial statements and, lagged by one quarter, and market prices. Since we are dealing with an international sample we use Datastream, a commonly used source for academics and practitioners. These variables are listed below.

IBES_TPP is the mean I/B/E/S target price for each firm divided by current price at the end of the previous quarter. This variable represents the consensus of forecasts made by each analyst covering a given firm; prior studies have compared individual analysts to consensus target price, earnings or recommendation score. *IBES_TPP* therefore provides a benchmark for *AVP* and *MVP*. Morningstar provides AVP to archives such as I/B/E/S; most other contributors are sell-side analysts. The fourth study discusses limitations of data such as I/B/E/S.

ROA is net income divided by total assets, both collected from Datastream at the end of the previous quarter. Since it can be calculated for most firms, it serves as a popular measure of quality (Piotroski, 2000). *ROA* is in the training set, i.e, it. is an input variable to ML.

BTM is the ratio of shareholders' equity to market capitalization at the end of the previous quarter. This measure of firm value is typically used in equity valuation and asset pricing models. It is included in Morningstar's training set alongside several other value multiples (e.g. earnings yield, sales yield, book yield, EV to market value). Since these are highly correlated, we opt for a simple structure by choosing only one valuation ratio, book to market, in our analysis in order to reduce multicollinearity.

FIRMSIZE is the natural log of firm capitalization measured by market capitalization of equity in US dollars.

Next, we include three variables based on market data only:

BETA is estimated from 60 monthly returns versus a market index, specifically S&P1500 for US firms or FTSE All Countries World ex-US Index for non-US firms. Since stock betas are notoriously unstable we multiply the estimate by 0.6 and add a constant 0.4 as a Bayesian adjustment. This approach is fairly standard in practice as it reduces the incidence of outliers, i.e., betas which are much higher or lower than one. The calculation is similar to the approach taken by data vendors such as Bloomberg.

VOLATILITY is the standard deviation of the prior 52 weekly (Datastream) total return observations.

VOLUME is the average number of shares traded in the past quarter obtained from Datastream. Volume is less typical in prior studies but is included in Morningstar's training set.

Finally we control for industry:

INDUSTRY is a dummy variable identifying the Global Industry Classification Standard (GICS) level 2 membership for each firm.

Our control variables replicate, as far as possible, the definitions provided by Morningstar. $FIRMSIZE_{i,t}$ is the log of US dollar market capitalization; $VOLATILITY_{i,t}$ is the standard deviation of the prior 52 weekly return observations.

Next, to test our hypotheses, we use cross-sectional regressions to examine the relationship between valuations made by analysts or ML and next quarter returns where the latter is the dependent variable. A panel regression approach would estimate the relationship over all quarters. We choose the Fama–MacBeth approach which has two steps. The first step is to estimate a cross-sectional regression in each period. The second step is to provide a summary of these regressions and this is done by averaging the beta coefficients. Since there may be autocorrelation in standard errors, we use the Newey–West approach (Newey and West, 1987) with two lags. **Equation 2** sets out our specification.

Equation 2:

$$RETURN_{i,t} = \alpha + \beta_1 AVP_{i,t} + \beta_2 MVP_{i,t} + \beta_3 RETURN_{i,t-1} + \beta_4\ IBES\_TPP_{i,t} + \beta_5 ROA_{i,t} + \beta_6 BTM_{i,t} + \beta_7 VOLATILITY_{,t} + \beta_8 LEVERAGE_{i,t} + \beta_9 BETA_{i,t} + \varepsilon_{i,t}$$

We run cross-sectional regressions using the Fama–MacBeth procedure. This design allows us to test all three hypotheses. If the coefficients on *AVP* or *MVP* are each positive, and statistically significant, we could conclude that both valuations contain investment value.

3.6.3 Data

Table 3.3 presents descriptive evidence on the characteristics of the firms in our sample. *AVP* and *MVP* each have means close to one. This confirms that the analysts are not biased towards high valuations. In contrast, the sell-side consensus, *IBES_TPP*, is substantially higher at 1.09 with a higher standard deviation.

AVP and *MVP* have close to zero correlation with *BTM.* This might be explained by Morningstar's standardized DCF approach. Surveys of sell-side analysts' practice report that DCF is less widely used than multiples (Brown et al., 2015). Closer correspondence can be observed between various risk measures. Volatility, beta and leverage all measure different aspects of risk but are positively related. As expected, *ROA* and *FIRMSIZE* are negatively related. None of the correlations are of concern when considering use in multivariate regression.

We also substitute *AVP* with *MVP* as the dependent variable and repeat the analysis, alternating AVP and *MVP* as independent variables to determine the relative importance of each as an empirical determinant of the other. Descriptive statistics of these variables are set out in Table 4.3 and correlations in Table 3.4.

Tables 3.5 and 3.6 set out the sectors and countries covered by analyst and the ML process. Many IRPs specialize in specific sectors. Table 4 shows that Morningstar's analysts cover firms in each sector without obvious concentration in any particular area. Morningstar has a long history of US equity research. The firm branched out into international coverage in 2012. Table 3.6 shows that over 53% of the sample is US. Outside the US, coverage is most concentrated in Australia, Canada, Hong Kong and the UK. The analysts do also cover stocks in many European and Asian markets and some emerging markets such as Brazil.

Table 3.3: Descriptive statistics

Variable definition can be found in table 3.2

Variable	N	Mean	Standard Deviation	Minimum	Maximum
$VOLATILITY_{t+1}$	19,839	0.03	0.13	-0.64	1.58
BETA (XUS)	19,384	0.85	0.25	0.40	1.60
BETA	19,839	0.99	0.33	0.01	2.37
VOLUME (M)	19,741	0.45	1.69	0.08	84.81
IBES_TPP	18,447	1.09	0.26	0.01	11.01
MVP	18,501	1.01	0.14	0.37	1.88
AVP	18,501	1.03	0.20	0.19	2.00
ROA	18,211	0.06	0.07	-0.74	0.58
BTM	18,501	0.53	0.42	-0.50	3.98
FIRM_SIZE	17,591	16.57	1.64	10.49	26.25
$VOLATILITY_{t-4}$	18,501	0.26	0.11	0.00	1.22
LEVERAGE	18,501	0.42	0.47	0.00	2.99

Table 3.4: Correlations

	RETURN t_{+1}	MVP	AVP	IBES _TPP	ROA	EQ/MV	EQUITY	FIRM SIZE	VOLAT ILITY$_{t-1}$	LEV’GE	BETA EX-US	BETA US
AVP	-0.303											
MVP	0.061	0.639										
IBES_TPP	0.021	0.146	0.173									
ROA	0.073	-0.211	-0.113	-0.115								
BTM	-0.001	0.000	0.001	-0.014	0.002							
EQUITY	0.002	0.030	0.013	-0.065	-0.002	0.039						
FIRM_SIZE	0.039	0.050	0.051	-0.006	0.084	0.071	0.267					
VOLATILITY$_{t-1}$	0.026	0.172	0.149	0.133	-0.257	0.013	0.003	-0.210				
LEVERAGE	-0.001	0.007	0.007	-0.011	-0.004	0.999	0.040	0.077	0.019			
BETA XUS	-0.009	0.186	0.132	0.134	-0.176	0.004	-0.014	-0.040	0.471	0.012		
BETA US	0.010	0.163	0.116	0.140	-0.146	0.003	-0.025	-0.053	0.419	0.010	0.942	
VOLUME	-0.001	0.131	0.087	0.013	-0.086	-0.001	0.043	0.281	0.103	0.012	0.125	0.088

Table 3.5: Industries

Industry	**N**	**%**
Automobiles and Components	457	2.31
Banks	1,146	5.8
Capital Goods	1,295	6.55
Commercial and Professional Services	434	2.2
Consumer Durables and Apparel	541	2.74
Consumer Services	600	3.04
Diversified Financials	1,047	5.3
Energy	1,459	7.38
Food and Staples Retailing	344	1.74
Food, Beverage and Tobacco	778	3.94
Health Care Equipment and Services	1,001	5.07
Household and Personal Products	252	1.28
Insurance	622	3.15
Materials	1,827	9.25
Media	440	2.23
Pharmaceuticals, Biotechnology and Life Science	1,052	5.32
Real Estate	901	4.56
Retailing	935	4.73
Semiconductors and Semiconductor Equipment	405	2.05
Software and Services	1,120	5.67
Technology Hardware and Equipment	440	2.23
Telecommunication Services	564	2.85
Transportation	806	4.08
Utilities	1,294	6.55
Total	19,760	100

Table 3.6: Geographic markets

Country	**N**	**%**
US	10,603	53.45
Australia	2,379	11.99
Canada	1,144	5.77
Hong Kong	1,102	5.55
United Kingdom	839	4.23
France	684	3.45
Germany	487	2.45
New Zealand	353	1.78
Switzerland	331	1.67
Japan	307	1.55
Netherlands	254	1.28
Singapore	219	1.10
China	178	0.90
Italy	157	0.79
Spain	151	0.76
Sweden	120	0.60
Denmark	118	0.59
Belgium	97	0.49
India	86	0.43
Norway	60	0.30
Finland	58	0.29
South Korea	33	0.17
South Africa	31	0.16
Portugal	20	0.10
Taiwan	17	0.09
Brazil	11	0.06
Total	19,839	100%

Tables 3.5 and 3.6 report the distribution of observations by industry (GICS level 2 sourced from Datastream) and geography (country of listing sourced from Morningstar)

3.7 RESULTS

Table 3.7 presents relationships between *AVP* (odd columns) and *MVP* (even columns) with key variables and controls. The first two columns show both valuations to be positively related to *IBES_TPP, FIRMSIZE, VOLUME* and *BTM*. Both sources of valuation are also positively related to volatility but not to beta. It may be that more volatile stocks provide opportunities for analysts to identify ideas to their stock-picking clients. Each valuation also has a negative relationship with *ROA*, but this is only statistically significant at 5% or 10% levels.

Table 3.7: Determinants of next quarter returns (Global)

	(1)	(2)	(3)	(4)
	AVP	*MVP*	*AVP*	*MVP*
$RETURN_{t-1}$	-0.0177	-0.167***	0.168***	-0.159***
	(-1.23)	(-10.88)	(15.34)	(-12.33)
IBES_TPP	0.151***	0.0527***	0.0911***	-0.0108
	(7.83)	(3.83)	(11.03)	(-1.60)
ROA	-0.0850*	-0.134**	0.0826**	-0.100**
	(-1.83)	(-2.30)	(2.17)	(-2.37)
FIRM_SIZE	0.0108***	0.00960***	-0.000823	0.00517**
	(3.39)	(3.06)	(-0.52)	(2.63)
VOLUME	1.49e-08***	1.28e-08***	-2.17e-10	6.73e-09***
	(4.15)	(4.48)	(-0.08)	(3.13)
BTM	-0.0191**	-0.00170	-0.0160***	0.00597***
	(-2.42)	(-0.52)	(-3.62)	(4.83)
BETA	0.00233	0.0200	-0.0166*	0.0188
	(0.12)	(1.12)	(-1.76)	(1.58)
VOLATILITY	0.207***	0.168**	0.0213	0.0852**
	(3.22)	(2.83)	(1.12)	(2.41)
LEVERAGE	0.0781***	0.0578***	0.0101**	0.0253***
	(5.33)	(5.10)	(2.17)	(4.03)
MVP			1.162***	
			(27.90)	
AVP				0.414***
				(52.97)
CONSTANT	0.655***	0.903***	-0.385***	0.632***
	(9.86)	(15.42)	(-10.02)	(18.12)
N	16,961	16,961	16,961	16,961
avg. R^2	0.1083	0.1886	0.4981	0.5464

t-statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Columns (1) and (3) report results for various specifications of equation 1A where analyst valuation (*AVP*) is the independent variable:

$AVP_{i,t} = \alpha + \beta_1 RETURN_{i,t-1} + \beta_1 IBES\_TPP_{i,t} + \beta_1 AVP_{i,t} + \beta_2 MVP_{i,t} + \beta_3 ROA_{i,t} + \beta_4 BTM_{i,t} + \beta_5 FIRMSIZE_{i,t} + \beta_6 VOLUME_{i,t} + \beta_7 LEVERAGE_{i,t} + \beta_8 BETA_{i,t} + \beta_9 VOLATILITY_{i,t} + \varepsilon_{i,}$

Where *AVP* is analyst valuation, *MVP* is ML valuation, *IBES_TPP* is the consensus target price collected from I/B/E/S, R*OA* is profitability measured using return on assets, *BTM* is firm value represented by book-to-market, volume is is the average number of shares traded in the past quarter, leverage is total debt to total assets, *BETA* and *VOLATILITY* are calculated using historic returns.

Columns (2) and (4) report results for various specifications of equation 1B:

$MVP_{i,t} = \alpha + \beta_1 RETURN_{i,t-1} + \beta_1 IBES\_TPP_{i,t} + \beta_1 AVP_{i,t} + \beta_2 MVP_{i,t} + \beta_3 ROA_{i,t} + \beta_4 BTM_{i,t} + \beta_5 FIRMSIZE_{i,t} + \beta_6 VOLUME_{i,t} + \beta_7 LEVERAGE_{i,t} + \beta_8 BETA_{i,t} + \beta_9 VOLATILITY_{i,t} + \varepsilon_{i,}$

The variable definitions as equation 1A apply

Where the determinants of *AVP* and *MVP* differ most is lagged returns, often referred to as momentum. Analyst valuations appear to be unrelated to past returns. Since Morningstar analysts use a standardized DCF model, and adjust valuations relatively infrequently, valuations are not closely related to historic returns. In contrast, *MVP* is automatically re-estimated on a daily basis. The input data used to train the algorithm draws on three main sources: reported financial statements which update two or four times each year, market data (price and volume) which update daily, and characteristics which seldom change (industry). The ML valuation is therefore much more sensitive to price changes than the analyst equivalent.

Columns 3 and 4 show the incremental effect of including *MVP* to explain *AVP* and vice versa. Each becomes the most significant determinant of the other and results in R-squared greater than 0.5, substantially higher than columns 1 and 2.

The significance of consensus target price persists in column 3, where *MVP* is used to explain *AVP.* This is not the case in column 4. Consensus information feeds into analyst valuations but is not used to train the ML algorithm. It seems that the independent analysts in our sample are suitably differentiated from the I/B/E/S consensus. This difference can at least partly be explained by the unbiased distribution of target prices compared to that of consensus; positive bias in the latter is well documented. Repeating the analysis on a subsample of US firms produces a similar pattern of determinants (not shown).

3.7.2 Information content of analysts' valuations

We hypothesize that independent analysts (hypothesis 1) and ML (hypothesis 2) each have some ability to predict returns. The first hypothesis is bold but informed by prior research by Allee et al. (2017), who find Morningstar's US analysts to have some predictive ability in an earlier period; the second has, to our knowledge, not been tested elsewhere. The outcome variable is next quarter return. We use the Fama–MacBeth procedure to estimate return using various combinations of *AVP*, *MVP* and control variables. The approach allows us to test both hypotheses.

Table 3.8: Predictors of next quarter returns (Global)

LHS:*RETURNi*$_{,t+1}$	1	2	3	4	5	6
AVP	0.0366***			0.262***	0.209***	0.197***
	(3.51)			(18.12)	(27.4)	(26.93)
MVP		-0.237***		-0.496***	-0.361***	-0.349***
		(10.44)		(19.60)	(23.16)	(26.87)
IBES_TPP			0.0127		0.0187***	0.0197***
			(1.67)		(4.56)	(6.35)
RETURN$_{t-1}$					0.157***	0.160***
					(18.41)	(21.98)
ROA						-0.00828
						(0.36)
BTM						-3.4E-05
						(0.04)
VOLATILITY						0.0229
						(1.25)
BETA						0.00592
						(0.58)
FIRM_SIZE						0.00268***
						(3.67)
CONSTANT	-0.00484	0.275***	0.0207***	0.263***	-0.0164	-0.0747***
	(0.60)	(13.16)	(3.19)	(16.08)	(0.99)	(3.72)
N	19,839	19,839	18,447	19,839	18,439	16,961
avg. R^2	0.0133	0.0928	0.0063	0.2028	0.3197	0.3751

t-statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.8 displays results for various specifications of equation 2

$$RETURN_{i,t} = \alpha + \beta_1 AVP_{i,t} + \beta_2 MVP_{i,t} + \beta_3 RETURN_{i,t-1} + \beta_4\ IBES\_TPP_{i,t} + \beta_5 ROA_{i,t} +$$

$$\beta_6 BTM_{i,t} + \beta_7 VOLATILITY_{,t} + \beta_{8i,t} + \beta_9 BETA_{i,t} + \varepsilon_{i,t}$$

R*OA* is profitability measured using return on assets, *BTM* is firm value represented by book-to-market, , leverage is total debt to total assets, *BETA* is the Bayesian adjusted beta = 0.4 + 0.6 * beta estimated from 5 prior years of monthly data vs S&P1500 for US firms and a World-ex US index for non-US firms.

Results for the full global sample are presented in Table 3.8. Columns 1 and 2 show *AVP* to be positively related to next quarter returns and *MVP* to be negatively related. The relationship with *MVP* is more significant and produces a higher R-squared of 0.09; although low, this is substantially higher than the R-squared of 0.01 attributed to *AVP* alone. Of particular note is the contrarian nature of *MVP*. How can this odd result be explained? Analysts are mindful that their investment management clients seek forward-looking (ex-ante) analysis. They change their valuation of any firm only occasionally, revisiting their DCF valuation when there has been an important change to the inputs or assumptions. The most obvious occasions would be the announcement of financial statements, guidance or industry news. The analyst may also revise valuations after conducting a detailed study of the firm, sometimes known in the industry as a "deep dive". If a stock price falls 5%, the analyst is unlikely to respond by lowering the valuation; such a move would be poorly received by clients who seek analysts who lead rather than lag..

ML estimates $AVP_{i,t+1}$ on a daily basis in order to obtain $MVP_{i,t}$. Although valuation changes infrequently, prices adjust each day. $MVP_{i,t}$ changes frequently in response to price changes. The result is therefore consistent with its greater sensitivity to price changes due to automated re-estimation. It should be possible to deal with this by including a momentum variable or, more directly, by amending the ML to target valuation without a price denominator.

Consensus target price, *IBES_TPP*, represents a widely available alternative to Morningstar's analyst and ML valuations and serves as a benchmark for our study. It does not add incremental information (column 3), nor is it significant when used as a single determinant of returns (not tabulated). Some prior studies have found some evidence that analyst target prices could help predict returns in the US (Brav and Lehavy, 2003) and Italy (Bonini et al., 2010) and in some other markets (Bilinski et al., 2012).

Next, we use both *AVP* and *MVP* to estimate returns (column 4) and find valuations are positively associated while ML valuations are contrarian. Despite this opposing direction of association, the R-squared increases substantially to 0.20.

Lagged return (column 5) is significant, confirming the importance of momentum in conjunction with *AVP* and *MVP*; R-squared jumps to 0.32. In columns 6 (*AVP*) and 7 (*MVP*), we introduce *ROA, VOLATILITY* and *BETA*, none of which are highly significant. *FIRMSIZE* is, however, significant. The R-squared of these models fail to match the simple combination of *AVP* and *MVP* in column 4. The most extensive model (rightmost column) achieves higher R-squared than the simple combination of *AVP, MVP* and lagged returns. When complemented by a momentum measure, analyst and ML valuations constitute a parsimonious means for informed stock selection. We repeat the analysis on US firms and find a very similar pattern of determinants (not tabulated).

The Fama–MacBeth procedure relies on cross-sectional regressions. We examined the underlying cross-sections to check that the overall results presented in this chapter are not driven by particularly strong relationships in certain quarters. This provides some comfort that average R-squareds are reliable. This review reduces the possibility that our results are influenced by a particularly strong relationship in some periods but not in others.

3.8 LIMITATIONS

While our study adds to the existing literature in several important ways, it has limitations. First, we use a single-firm study. This is necessary due to data limitations. Morningstar is, to our knowledge, the only firm to use ML to imitate analyst valuations and to have made these widely available. Second, since our analysis is quarterly, ML valuations are fresh and analyst valuation changes may be stale. Since analyst valuations appear to be more informative this does not appear to be a major problem. Third, we are unable to identify individual analysts (the firm does not provide a mapping of historic coverage). Analyst dummy variables are likely to be correlated to sector, which we include. Fourth, we have no descriptive information about analysts. The latter issue could partially be addressed using public online information such as Linkedin. Finally, Morningstar do not state how they address countries and regions. No mention of a country variable is given.

Further research might train the ML based on different inputs, for example including price momentum, and considering alternative dependent variables such as valuation, or the ex-ante return implied by the valuation (rather than value to price). It would also be interesting to test whether analysts learn from ML. This could be done by researchers who have built their own ML models.

A final limitation is that if analyst's predictions may not be especially valuable to investors. Imam and Spence (2016) document an emerging line of research on the broader role of analysts. Summary outputs, in our case valuations, are now thought to be less important than other aspects of analysts' work. In the following chapter we complement valuations with other outputs: risk assessments. Thus I take a step towards a more holistic examination of ML applications in investment research.

3.9 CONCLUSIONS

Firms are employing ML to produce investment research. This study contributes evidence from an early adopter of ML in investment research. Using the Fama–MacBeth procedure, we compare the informativeness of valuations made by independent analysts and by an ML algorithm designed to mimic their outputs. Our sample is particularly useful because the analysts produce an unbiased distribution of forecasts.

Analysts' ability to predict next quarter returns is insufficient for investment purposes. In contrast, ML valuations are significant but contrarian. Automated daily re-estimation makes ML valuations substantially more sensitive to stock returns. When used together, analyst and ML valuations explain some 20% of variation in next quarter returns, which increases to 32% when combined with price momentum. In contrast, I/B/E/S consensus target price is unrelated to future returns. We confirm our findings on a US subsample representing 62% of our dataset.

International studies of target prices draw upon archives, datasets which are dominated by brokerage analysts. This chapter provides evidence from independent analysts, an important, if less documented, category of agent.

For investors, the key implication of our study is that a combination of analyst and ML might be superior to using either input alone. For analysts in our sample firm, the first lessons would be to consider the importance of momentum, and sentiment more generally, in reaching their valuation, and the second lesson would be to consider using ML to identify improvements. For ML developers, the selection of input variables, again including price momentum, is also crucial.

Appendix – Key papers on analyst target prices and non-brokerage analysts

Panel A: Target price studies

Paper	Data source	Sample	Main findings
Brav and Lehavy (2003)	First Call	1997-1999	Target price revisions generate short term price reaction which is not explained by changes to recommendation and earnings forecasts, and longer term (up to six month) drift in the direction of the target price revision.
Asquith, Mikhail and Au (2005)	Investext, IBES, Zacks	1997-1999	Target prices supply incremental information which is not fully encapsulated in earnings forecasts or recommendations.
Bonini et al (2010)	Hand collected reports, Italy	2000-2006	Analysts exhibit ability to forecast prices. Forecast error increases with growth predictions, size, negative earnings, research intensity (analysts covering more stocks) and bull markets.
Demirakos, Strong and Walker (2010)	Hand collected reports	2002-2004	Four methods are used to test the accuracy of analysts' target prices. Multiples (PE) outperform discounted cash flow (DCF) models where the target price is met during the 12-month forecast horizon but not in other tests, e.g., where target price is not met within 12 months.
Gleason, Johnson and Li (2013)	First Call	1997-2003	Target price accuracy is higher for analysts who use rigorous valuation methods such as residual income or discounted cash flow, and who have greater accuracy in predicting earnings.
Bilinski, Lyssimachou and Walker (2013)	I/B/E/S 16 Countries	2002-2009	Significant variation in average target price accuracy across countries is explained by accounting disclosure quality, the legal system type, cultural characteristics and IFRS. Some analysts demonstrate persistent (differential) ability to forecast target price.
Bradshaw, Brown and Huang (2013)	I/B/E/S	2000-2009	Some analysts demonstrate persistent ability to forecast target price; while these results are statistically significant the economic significance is marginal and a strong optimism bias persists. Main contribution is the large sample compared to Brav and Lehavy (2003).

Bradshaw, Huang and Tan (2014)	I/B/E/S/ Nelsons 44 Countries	2002-2011	'Target price optimism is positively associated with proxies for analysts' conflicts of interest, but negatively associated with country-level institutional infrastructure as characterised by strong investor protection, effective legal enforcement and transparent financial environment'
Dechow and You (2017)	I/B/E/S	1999-2012	Target price errors can be explained by three sources: firm risk characteristics, errors in forecasting fundamentals and biases related to analyst incentives. Risk is the most important source.
Ho, Strong and Walker (2018)	I/B/E/S UK	2003-2014	Controlling for earnings forecast and recommendation revisions, target price revisions are associated with recent market returns, excess stock returns, and other analysts' target price revisions. Target price revisions are more sensitive to negative than to positive excess stock returns.

Notes: all studies examine sell-side analysts.

Panel B: Independent research and the effects of the Global Settlement

Paper	Analyst type(s)	Unit of analysis	Sample	
Kadan et al. (2009)	SS/IRP	Recommendations	2000-2004	Diff-diff. study comparing periods before/after Global Settlement. Brokers shift to 3-point (B/H/S) rather than 5-point. 12% fall in average #stocks covered. Recommendations became less informative based on 3-day price reaction.
Barber et al. (2006)	SS	Recommendations	1996-2003	After Global Settlement: Fewer buy recommendations. Brokers with less uneven distribution of forecasts tend to be more accurate. IRPs are analysed separately in Barber et al. (2007).
Barber et al. (2007)	SS/IRP	Recommendations	1996-2003	Portfolio tests show unaffiliated analysts' buy recommendations outperform investment bank (affiliated) comparators; IB hold/sell recommendations outperform. IRPs and unaffiliated are together labelled 'Independent'.
Jacob and Rock (2008)	SS/IRP	Earnings	1995-2003	IB analysts are more accurate than IRPs. IB cross subsidise -> superior resources, pay and access to (some) IB knowledge. 'Quality of research' explicitly proxied by forecast accuracy.
Clarke et al. (2009)	SS/IRP	Recommendations	2000-2007	After GS: Fewer Buy recommendations, downgrades and IRP recommendations became less informative. Recommendations became less informative based on price reaction.
Buslepp et al. (2014)	SS/IRP	Recommendations		IRPs funded by GS are less accurate than brokerage analysts and other IRPs and this result can be attributed to

				the relative experience of IRPs.
Allee et al. (2017) Working Paper	SS/IRP	Earnings and target prices		A single firm study of Morningstar analysts provides evidence that IRPs earnings forecasts are more accurate, especially for stocks with high momentum and high valuations; the improvement may be due to valuation approach and more modest long-term growth estimates.

Panel C: Single firm studies: comparisons of buy-side and sell-side analysts' forecasts

Paper	Analyst type(s)	Unit of analysis	Sample	
Groysberg et al (2008)	SS/BS	Earnings	Single BS 1997-2004	BS analysts' earnings forecasts were more optimistic and less accurate than sell side forecasts. This may be because the BS firm has different forecasting horizon or because it retails analysts who are persistently inaccurate.
Rebello Wei (2014)	BS	Return forecasts	Single BS 1994-2008	BS research influences the holdings of portfolio managers, especially for stocks with little sell side coverage. BS analyst's skill levels persist through time and overall have some predictive power.
Cheng et al (2006)	BS	Fund holdings	Multiple BS 2000-	Empirical tests of a model which combines biased SS analyst research with unbiased BS analyst recommendations. Fund managers have differential use of BS vs SS research as evidenced by a market survey conducted in 2000.
Frey and Herbst (2014)	BS	Recommendations and fund holdings		BS analysts' recommendations influence decisions of portfolio managers within the firm

				more than SS analysts' recommendations. Trades based on BS analyst recommendations improve fund performance.
Cici and Rosenfeld (2016)	BS	Fund holdings	Single BS	A study of funds managed by BS analysts within one BS firm. BS analysts demonstrate ability to generate alpha for their own fund and also have some influence on the holdings and performance of funds managed by their colleagues.
Crawford et al (2017)	BS	Recommendations	Multiple BS 2010	Buy side recommendations posted on Sumzero (a private social network) are informative as evidences by short term reaction and longer term drift. Contrarian recommendations tend to produce larger and more significant returns.

CHAPTER 4

The potential for machine learning based investment analysis: analyst versus machine learning based risk assessments

'No human can beat a computer at chess. And no computer is better at chess than a human supported by a computer.'

Lopez de Prado (2018, 15)

4.1 INTRODUCTION

Machine learning (hereafter ML) algorithms could become ubiquitous in investment analysis. Certain tasks performed by analysts, such as risk assessments, could be systematically generated and calibrated on a continuous basis. Successful application could substantially extend coverage, reduce costs, mitigate biases found in human decision-making and, possibly, improve the quality of predictions. Early signs of adoption, which could change the role of both analyst and investment manager, are already appearing. Leading buy-side and sell-side firms have recently unveiled prodigious capabilities in ML, but the potential for change surely hinges on whether ML can match or even surpass analysts in predictive tasks. We contrast the ability of analysts and ML in quantifying the distribution of possible valuation outcomes.

The prospect of wider coverage with frequent, cheaper and less biased valuations makes ML appealing. Yet, while ML has advantages in speed, scale and objectivity, human fundamental analysts, powered by expertise and intuition, can process unstructured data to identify and evaluate news, e.g. managerial changes, contracts won or controversies exposed. A systematic approach to ML may struggle to capture subtle contextual information. The potential for ML therefore depends upon its effectiveness when contrasted with alternative approaches, most obviously existing human approaches.

Our setting allows us to examine the potential for ML based investment research. We use risk measures derived from data provided by an early adopter of ML, an independent research organisation with over 100 analysts covering some 1,500 firms globally. In 2012 the firm launched ML based trading recommendations, valuations and risk assessments for more than 20,000 firms.

We use subsequent return volatility to measure the effectiveness of risk assessments. Prior evidence suggests that financial analysts can assess risk with substantially more success than predicting returns (Joos et al. 2016). In an efficient market, only exceptional individuals will consistently beat the market. There is no equivalent competitive element to risk assessment. For most stocks, risk in the next period is strongly associated with risk in the last (Poon and Granger 2003, 2005). Thus, whether analysts or ML can provide forecasts that have predictive ability is not contentious: it would be surprising if they did not. Analysts should be able to make accurate risk assessments. We compare analyst and ML risk assessments.

4.1.1 Analyst risk assessment

Analysts in our sample firm allocate firms to a low, medium, high or very high risk category. Each category is mapped to the distribution around each analyst's valuation estimate. Rather than allowing analysts to produce refined, granular risk assessments, Morningstar opts for simplicity by standardising the range between the boundaries for each category. The his is done by calculating the range between the upper ("bull") valuation and the lower ("bear") valuation, divided by the valuation estimate. For the lowest risk category, the standardised range is 44%. Morningstar analysts will assign a variety of recommendations for these firms, but all firms within the category have the same standardised range, i.e., the same risk assessment when expressed as a percentage range, in this case meaning that the analysts expect the value to be within 22% above or below the valuation. The standardised ranges for the medium, high and very high categories are 63%, 88% and 111%, respectively. The mean range for our sample is 76%, i.e. between medium and high risk. Approximately 5% of firms are designated low, 47% medium, 38% high and 10% very high risk. The use of categories is similar to the approach used in Liu et al. (2007).

The analysts do not have the opportunity to make more refined risk assessments, e.g., by setting other ranges or by allowing a wider range either above or below their valuation. Their only alternative would be to choose a different category. This risk assessment is coupled with the analysts' valuation to price computation to determine the investment recommendations: the riskier the stock, the more extreme the valuation to price required to justify a buy or sell recommendation. Risk assessments are therefore directly plugged into the firm's valuation process.

4.1.2 ML risk assessment

The firm's ML approach is based on the random forest method (Breiman 2001). The intention is that ML will replicate each analysts' output. ML produces 500 valuation estimates, each an iteration of the model estimated on a random subsample of the available data. The output is then summarised as a mean, the ML valuation and MRisk, which is the interquartile range of 500 valuation estimates. The research provider specifically describes this as comparable to the analyst risk assessment.

Our search for prior studies of ML in investment research therefore reveals that while practitioners invest substantial sums to forge ahead with ML applications, there is an important gap in the literature. Fortunately, our dataset provides a near-ideal setting. Our sample, MRisk, has a mean of 14% and lower and upper quartiles of 9% and 17% respectively. ARisk is therefore more than five times higher in mean score (0.76) than quantitative score (0.14).

4.1.3 Future volatility

Our outcome variable is return volatility for each security, measured as the annualized standard deviation of weekly returns for the forthcoming quarter. Volatility calculations using weekly data are helpful for our global sample as they are less susceptible to thin trading or price-timing issues. For example, news may arrive after Asian markets have closed but during US trading. Liu et al. (2007) also select volatility as their outcome variable, in their case using the log of daily volatility for their US sample. They also present a sensitivity analysis to demonstrate that their results are not affected by the precise calculation of volatility, including various calculation windows. We choose volatility in the subsequent quarter as it allows us to analyse non-overlapping data and therefore reduces the chance of incorrect inference.

Some papers choose valuation error rather than volatility as the outcome variable. Joos et al. (2016) examine the valuation error (the difference between analysts' valuation estimate and the share price one year later). Alternative measures could include the range of prices in a subsequent period and non-symmetrical measures such as semi-deviation. The latter is designed to separate downside (possible loss) from upside (possible gain, i.e., opportunity).

Although designed to be comparable, the two risk assessments (ARisk and MRisk) and the outcome (volatility) are measured differently. The mean ARisk and MRisk for our main sample is 0.142 and 0.767 respectively, whereas the outcome variable has a mean of 0.239. This difference complicates the interpretation of coefficients. We mitigate this problem by using Fama-MacBeth regressions to derive a forecast of volatility from the data at each quarter and then assessing the explanatory power of the derived forecast in the following quarter. Our three derived forecasts, one based on the ARisk measure, one on the MRisk and a third estimated using the financial characteristics of the firms, all have the same mean, 0.239, close to that of the outcome volatility. Each is explained in turn in the following subsections.

4.1.4 Derived forecasts

We estimate the relationship between the risk indicators at the end of each calendar quarter, labelled *t,* and volatility in the subsequent quarter, denoted *t+1*. The estimated relationship from this cross-sectional regression allows us to predict volatility at *t+1* based on the risk indicators at *t*. In other words, we are estimating volatility in each subsequent quarter. Using Fama-MacBeth cross-sectional regressions, we evaluate the effectiveness of the forecasts by estimating the explanatory power of the forecast in each quarter and the average across the 27 quarters in our seven-year sample period. (One quarter is necessarily dropped as the risk measures at the end of 2018 would need to be assessed against volatility in the first quarter of 2019.) We have named the analyst-derived forecast (ADFor) and the machine-derived forecast (MDFor). This approach gives us a means to deal with risk assessments on different scales and is an innovation not included in previous studies, which typically estimate the relationship across a panel of firms and time.

The third forecast we use follows a parallel approach, except that we based the forecast on the financial characteristics of the firms: the volatility in the previous quarter, the market capitalization of the firm, the firm's beta and the dispersion of analyst valuations available on I/B/E/S. (We also use these variables as control variables because they represent important risk characteristics and are similar to those used in Joos et al. (2016)). This forecast is a simple model of the forecast an investor could access should they not have access to the analyst or machine risk measures. We name this the financial-derived forecast (FDFor).

To provide a brief preview of the results, both analyst and ML risk assessments contain information on risk which is incremental to that contained in the typical risk model factors. Each can be improved by using information contained in the other. These results hold broadly across time and countries. However, the different constituents of the sample present different challenges. To examine this, we split the sample into US and non-US segments. We also compare common law countries, where we argue that the information available to investors may be more complete, and code law countries, where the information environment may be less complete. We also divide the sample into cases where I/B/E/S collates valuations and forecasts, and those without I/B/E/S coverage. The presence of an I/B/E/S following indicates that at least one analyst is monitoring the firm and disseminating information to investors.

Analyst- and machine-derived forecasts are both less effective indicators of volatility in code law countries than where common law is applied. The decline in analysts' predictive ability across this divide is such that the ML predictions marginally outperform those from analysts for the code law sample. This indicates a potential advantage beyond the ability of machine-based techniques to produce more frequent, cost-effective and comprehensive coverage than is feasible with analysts: for our sample, the ML approach handles unfamiliarity and limited informational environments better than analysts. However, if this were a general result, we would expect analysts to cope less well with firms not covered by other analysts (i.e. with no I/B/E/S following). Our results show no appreciable difference between firms with and without I/B/E/S coverage.

As we study analysts from an independent research provider, the adverse incentives faced by sell-side analysts are removed. This bias is most obviously demonstrated by the valuation to price assessment. For example, Joos and Piotroski (2017) report a mean valuation to price of 1.159 for their (2007-2012) Morgan Stanley sample, with only 22% of cases below one. The mean valuation to price for our sample is 1.055, with 45% below one. Yet, despite the absence of typical sell-side incentives, our tests still demonstrate mild bias. We find that analyst risk assessments are relatively narrow for companies for which they issue buy recommendations. Since the analysts in our sample do not face brokerage incentives, we infer that this bias is behavioural. The result is consistent with self-attribution of knowledge and overconfidence: experts have often been found to underestimate the range around their forecasts (Barberis and Thaler, 2003). We find that analysts do not impart bias on the ML risk assessments.

The main contribution of this paper is to provide early evidence of the prospects for ML by evaluating the effectiveness of ML in performing tasks usually done by experts. The experts in our study are analysts at an independent research firm; biases in their outputs cannot be the result of sell-side incentives and therefore can be attributed to human behaviour. Rather than relying on stylized models, or attempting to present the best possible predictor, our data comes from a well-established, commercially available source. Such data is typically inaccessible (Brooks et al. 2019) Additionally, our work is related to recent research on the second moment of analysts' predictions and the expanding independent category of investment analysts (see for example Joos et al. 2016). Finally, we present global results and therefore build on the minority of studies featuring global or international findings and analysts' work.

The study has a number of limitations. As with most studies of risk assessment, we use a single-firm setting. We present evidence that analyst and ML outputs are credible and widely available, and we make the case that the study of an early adopter is useful given the increasing interest in ML investment applications. Our sample firm uses supervised ML to mimic analyst outputs from a narrow range of input variables.

These variables may become more or less important over time. Other supervised ML algorithms might use wider sets of input data. Other supervised learning approaches, such as deep learning and neural networks, might also be used and may be more or less effective (Gu et al. 2018). Unsupervised learning approaches are also popular and available to discretionary investment managers, i.e. where the ultimate investment decision is subjective rather than systematic (Lopez de Prado 2018). Examples include algorithms which establish sentiment scores from unstructured text. We expect a surge in the study of ML in investment management and research.

4.2 PRIOR RESEARCH

Stock valuations, or target prices, provide a more informative summary than recommendations or earnings forecasts (Bonini et al. 2010; Bilinski et al. 2012; Bradshaw et al. 2012). A recent, systematic review of the research concerning analysts in accounting, finance and management journals (Spence et al. 2019) reveals that two thirds of papers study earnings or recommendations, compared to only 10% examining other forecasts, of which target prices are only one type.

More recently, analysts at some firms have included summary information about the distribution around their target price. These risk assessments, sometimes known as bull/bear analysis (BBA), usually take the form of alternative target prices for positive and negative scenarios. Several recent papers relate risk assessments to outcome variables which represent fundamental risk, such as future share price volatility (Liu et al. 2007) and absolute pricing error (Joos et al. 2016). Further, Liu et al. (2012) report that the publication of analysts' risk assessments affects share prices and investment returns, and Joos and Piotroski (2017) show how the relationship between the distribution of scenario (BBA) valuations and the analysts' target price can be used to discriminate between more and less accurate investment predictions. Hashim and Strong (2015) also find that target prices accompanied by risk assessments outperform those without. Hashim and Strong (2015) use a wide-ranging sample of reports on US firms, while Liu et al. (2012) and Joos and Piotroski (2017) use single firm settings; all use risk assessments provided by sell-side analysts, with their well-established propensity for biases 2, and all concentrate on US firms. To our knowledge, technological innovation, automation or machine learning have not previously been investigated in this strand of the literature.

The Global Settlement and associated legislation placed restrictions on investment bank analysts and prompted a flurry of studies comparing investment bank and non-investment bank analysts. Prior studies tend to group independent analysts with 'unaffiliated' peers, who work at non-investment bank brokerage firms, and refer to this category as 'independent'. Haig and Rees (2016) provide a discussion of the regulatory changes between the Global Settlement and current European regulation (MiFID II); they document that investment managers are becoming increasingly interested in independent investment research. MiFID II addresses the conflicts faced by brokers regardless of their affiliation to investment banks by 'unbundling' execution and research commissions, which means that the method used to pay brokers is now the same as that used to pay independent research firms. Studies of the characteristics of independent analysts are therefore likely to be valuable. Our analysts therefore offers a preview of what we might expect to find in future.

4.3 HYPOTHESIS DEVELOPMENT

Analysts, like many experts, are expensive to employ and are subject to human error. They make mistakes, such as those identified by Barberis and Thaler (2003). The possibility of using systematic processes to augment or replace expert decision-making is therefore of interest. Simple static models have been shown to be more accurate than experts in domains such as psychological diagnosis (Meehl 1956), political forecasting (Tetlock 2005) and systematic, quantitative investment (Fabozzi et al. 2010). Systematic investment models have tended to be static and to require analysts' skills and experience to determine when and how they should be re-estimated in order to deal with living, evolving financial markets. Models which not only update data but also re-estimate the relationship between variables at short time intervals, e.g. daily or intraday, could replicate or even surpass experts. Developments in ML have led to the introduction of dynamic models which learn and update without human intervention.

Risk assessment differs from volatility prediction. The former is a means for analysts to convey their expectations regarding the distribution around the target price; it is an indication of risk rather than a precise estimate of ex ante return variance (Joos et al. 2016). The latter, volatility prediction, has various purposes, including options trading and risk management. The substantial literature on volatility prediction, reviewed by Poon and Granger (2003, 2005), 'reflects the importance of volatility in investment, security valuation, risk management, and monetary policy making' (Poon and Granger 2003, 478). Implied volatility offers an appealing alternative estimate of risk, a forward-looking consensus estimate derived from option prices. Recent papers, such as those by Baltussen et al. (2012) and Szado et al. (2018), review the potential for implied volatility to be incorporated in stock selection strategies. All these studies use US-implied volatility as a predictor; many use data from OptionMetrics, who provide US data on several thousand US companies. Even so, studies such as Szado (2018) restrict their analysis to S&P500 companies due to liquidity constraints. Outside the US, exchanges make single stock options available foe a narrow selection of the largest listed companies. In the UK, exchange traded options exist for around 120 companies; in France and Switzerland, options are available only for constituents of the large cap indices and a few additional large companies. Additionally, implied volatility is not reliable for illiquid options in these large companies. Our own discussions with option-savvy practitioners confirm that it is impractical to use implied volatility as an input to broad market equity strategies in non-US markets. Option implied volatility is only available for liquid single stock options, which makes it much less applicable for many companies in our sample. Since historic volatility is generally available, we include it as an explanatory variable in order to understand whether analysts or ML can provide incremental information in their risk assessments.

4.3.1 Research Question 1: Can ML mimic analyst risk assessments?

The ML algorithm learns from and attempts to mimic analysts' target prices. If successful, we expect analyst and ML risk assessments to be related:

> ***Hypothesis 1 (H1): Analyst risk assessments vary with ML risk assessments.***

This hypothesis is a basic preliminary for our analysis. We test this hypothesis by comparing the determinants of analysts' risk assessments.

4.3.2 Research Question 2: Are analyst or ML risk assessments more informative?

Investors will surely be interested not only in the extent to which ML can mimic analyst risk assessments but also in the information they provide.

Prior studies of investors' geographic proximity to investee companies have often found a home bias. This is usually explained either by the familiarity hypothesis or the information asymmetry hypothesis. Familiarity may be due to cultural, linguistic or emotional factors (Grinblatt and Keloharju 2000) or behavioural biases, such as overconfidence and self-attribution of knowledge (Barberis and Thaler 2003). Analysts might therefore be likely to place too narrow a range around their target prices for companies with which they are familiar, such as those in their home market.

Information asymmetry explanations for proximity effects are based on the idea that analysts who can make 'house calls rather than conference calls' (Malloy 2005, 1) are in a better position to find and interpret news which could affect local companies. Investment managers have been shown to generate higher returns on local rather than distant firms. Coval and Moskowitz (2001) attribute this to the value of local contextual information, such as face-to-face meetings with management. Similarly, analysts meet CEOs face-to-face and survey a firm's operations directly, and use this private information to make more accurate predictions (Malloy 2005).

Since our data provider has a longer history of covering US companies than international firms, and the majority of analysts are based in the US or other common law countries, we can expect analysts to be more accurate in calibrating valuations for equities listed in the US and other common law countries, because of proximity to analysts. Few analysts in our sample are based in code law countries:

> ***H2a: Analyst risk assessments will be most informative for US stocks.***
> ***H2b: Analyst risk assessments will be most informative for common law stocks.***

4.3.3 Research Question 3: Do independent analysts produce unbiased forecasts of risk?

Analysts have been found to exhibit optimism bias due to behaviour and incentives. Sell-side research firms typically have more buys than sells. By contrast, Morningstar analyst recommendations follow a near symmetrical, bell-shaped distribution. We can therefore determine whether independent analysts suffer from the behavioural biases found in other samples (Groysberg et al. 2008). Our sample is unlike that of prior studies which merge independent analysts with their peers at non-investment brokerages, and it provides an unusual opportunity to study truly independent analysts.

If independent analysts place a narrower range around buy-rated stocks than sell-rated stocks, this would be evidence of self-attribution of knowledge, i.e.

overconfidence in having picked a winner. If analysts forecast a narrower spread for US stocks than international stocks, after adjusting for risk, this would be evidence of familiarity bias. If analysts do not suffer from self-attribution bias, we expect analyst and ML spread to be similar for buy/neutral/sell-rated stocks.

> ***H3a: Analyst risk assessments underestimate risk for buy-rated stocks (self-attribution bias).***
>
> ***H3b: ML risk assessments do not vary between (implied) recommendation category.***

4.3.4 Research Question 4: How can investors best use analyst and ML risk assessments?

This paper could be framed as a contest between human and artificial intelligence. It is, however, possible that the best risk predictions can be made by a combination of mind and machine. In the absence of perfect correspondence between the two sources, it may still be possible for investors to combine the two valuations to their advantage. Taking stock of recent evidence on the use of ML to complement analysts' earnings forecasts (Ball and Ghysels 2017), we have at least some basis to expect that a hybrid approach may be optimal. Our final hypothesis is:

> ***H4: A combination of ML and analyst makes superior predictions (than either source in isolation).***

Our large sample of ML-only predictions allows us to perform further tests on this hypothesis.

4.3.4 SETTING, DATA AND METHOD

Our data comes from a relatively large independent provider of equity research. With around 100 analysts covering securities issued by 1,500 companies, its scale is comparable to a large brokerage firm (although smaller than the very largest global investment banks) but without the incentives faced by sell-side analysts (Barber et al. 2007). Analyst data is available from 2002; over time, the coverage expanded from the original US sample to include all sectors. Each analyst covers, on average, 15 firms and typically holds similar credentials to sell-side analysts (Kang et al. 2018). Morningstar's equity research reports and outputs resemble those of a brokerage firm, and they provide two risk assessments. The first is an analyst risk classification into low, medium, high or very high, where each category is also allocated a score reflecting the interquartile range of expected investment outcomes. The second, ML-derived, is the interquartile range of valuations derived from 500 iterations of their random forest model.

We also analyse a sample of firms for which Morningstar provides ML data without accompanying analyst output. This is a much bigger sample which reflects the wider coverage using the ML approach. The firms in this extension sample are often drawn from less developed or developing economies where the information environment might be expected to be shallower than in countries with long-established stock markets. When we match ML and analyst data in our comparison sample, we include 969 target firms. When we analyse firms with ML data but no matched analyst data, we have 11,164 firms (extension sample).

Although the ML approach used is proprietary, the firm does provide documentation, and their research team assisted us in answering further questions. The input factors resemble those used in established stock selection tools. The firm makes their outputs widely available; we estimate that some 3,000 buy-side firms have access. It is reasonable to assume that other firms have developed comparable and possibly superior models, but these are much less widely available. Our setting therefore provides a unique opportunity to evaluate the effectiveness of a working ML process rather than a prototype model.

4.4.1 Research method

Liu et al. (2007) and Joos et al. (2016) first model the determinants of the analyst and ML risk assessments and future volatility. Although this analysis is descriptive, if future volatility is associated with a particular variable, and that variable is not associated with the risk measure, it would suggest that the risk measure neglects relevant information. To ensure consistency with previous studies, we have reported determinants by including controls for past volatility, the log of market capitalization and beta. We also control for the dispersion of I/B/E/S target prices, since this measures the variety of valuations made by analysts.

We also considered including option-implied volatility as a viable forecast of volatility. Using implied volatility would be feasible for most US stocks, and a subset of the larger non-US firms, and we might expect analysts to refer to option prices in forming their expectations. However, deriving viable implied volatilities is not straightforward when the liquidity, moneyness, and horizon of options differ. Poon and Granger (2005) point out that implied volatility forecasts work well if calibrated by historical volatility, and that historical volatility is also an effective model when used alone. Our controls include historical volatility, and we experimented with the volatility index (US VIX and CBOE's ex-US equivalent) to incorporate changing market-wide expectations, but this, while statistically significant, had little impact on our research question. We have also included the dispersion (coefficient of variation) of target prices available on I/B/E/S to represent uncertainty. Our goal is not to develop the best possible risk prediction model but to contrast the effectiveness of analysts and ML risk metrics. The control variables are used to ensure that the risk metrics provide information beyond what is readily available to investors.

4.4.2 Information content of analyst and ML risk assessments

To identify the relative information content of the two risk assessments, we estimate the results using the analyst- (*ADFor*), ML- (*MDFor*) and financials- (*FDFor*) derived forecasts separately and then together, with and without the control variables. Equation 1, estimated using Fama-MacBeth estimation, is:

$$VOLATILITY_{i,t+1} = \alpha + \beta_1 XFor_{i,t} + \beta_2 VOLATILITY_{i,t-4} + \beta_3 VOLATILITY_{i,t-1} + \beta_4 FIRMSIZE_{i,t} + \beta_5 BETA_{i,t} + \beta_6 IBES\_CV_{i,t} + \varepsilon_{i,t}$$

Where $XFor_{i,t}$ is $ADFor_{i,t}$ or $MDFor_{i,t}$ or $FDFor_{i,t}$, respectively forecasts of forthcoming volatility derived from analysts, ML or financial-based risk assessments at time *t*, and the relationship between those variables at *t-1* and volatility at *t*, $VOLATILITY_{i,t+1}$ is volatility for the forthcoming quarter, $VOLATILITY_{i,t-1}$ and $VOLATILITY_{i,t-4}$ are the last quarter's and last year's volatility, $FIRMSIZE_{i,t}$ is the log of market capitalization in US dollars, $BETA_{i,t}$ is the Bayesian-adjusted beta estimated in the prior 36 months using either the US market index for US firms or the FTSE world (ex-US) index for non-US firms, and $IBES\_CV_{i,t}$ is the coefficient of variation of the I/B/E/S reported target prices. Definitions of all variables are provided in Table 3.1.

This model is similar to approaches used in previous studies, save that we omit accounting-based variables such as book-to-price and return on assets. In initial analysis we find that these variables can be statistically significant but do not affect our conclusions. We have also included the coefficient of variation of I/B/E/S target prices as an alternative measure of the second moment of valuation, which is also based on a distribution of valuation estimates, in this case from the cross section of analysts covering each given firm. We find this variable to be an effective indicator of future volatility. Finally, we have chosen to use Fama-MacBeth cross-sectional estimation (Fama and MacBeth 1973). Previous research has typically used panel data estimation, and we find the results are similar whether we use panel or Fama-MacBeth estimation. In practice, a user of analyst output will have to predict the next period's volatility using currently available indicators and experience of the relationship between those indicators and volatility in prior periods. By including forecasts derived from the previous quarter's relationship between the indicators and volatility, and evaluating the performance of those derived forecasts quarter by quarter, we closely simulate the decision-making circumstances of a user of investment analysis – whether performed by analysts or ML.

A feature of our sample is that analyst and ML valuations are updated at different intervals. Analysts tend to revise their targets intermittently, typically 3-4 times per year. ML is re-estimated daily. Quarterly cross sections provide a balance between these two frequencies. Since next-quarter returns are the key dependent variable in our analysis, this also means that there are no overlaps in our dependent variable.

Fama-MacBeth regressions have other advantages. Petersen (2009) shows that the approach is appropriate where residuals are correlated across firms within any period, known as a time effect. In the same paper, Petersen reports that this attribute has often been used to claim that Fama-MacBeth standard errors are unbiased. Unfortunately, this need not be the case, as there may be a 'firm effect' where residuals for any one firm are correlated through time. The two effects can be identified in a dataset using a combination of clustering approaches. We can compare the significance of each coefficient produced using clustering by firm and time (two-way clustering) with the Fama-MacBeth regressions, where the latter incorporates a robust approach to deal with autocorrelation (Newey and West, 1987). What we are looking for is a small difference between these two robust approaches. A further practical test can be obtained from our Fama MacBeth analysis. Rather than simply relying on the average coefficient, we can perform a count of the direction and significance of each coefficient in each regression. If a coefficient tends to be, for example, positive and significant in many of the cross sections, then we can have greater confidence in our inference.

4.4.3 Data

In our dataset, we have two different indicators of risk from the same supplier. The first is a traditional risk classification into low, medium, high and very high (plus extreme, but we have not identified any such case in our sample). This is notionally built on BBA with a 25% chance of exceeding each boundary, and the spread between the two can be interpreted as the interquartile range. This range is used to classify each firm into one of five categories, with 5% classified as low risk, 48% medium, 38% high and 9% very high. The range between bear, designated 'consider buy', and bull, 'consider sell', divided by the average of the two, is 44%, 63%, 88% and 111%, respectively. We use these to create ARisk in the results presented. Morningstar has chosen the blunt simplicity of ordinal categories rather than allowing analysts to assign precise ranges.

The second Morningstar variable is a measure of dispersion within the valuation model, calculated as the interquartile range from 500 trees in the random forest. It therefore reflects uncertainty but is not explicitly a prediction of expected risk. In essence, a model is estimated which relates the predicted variable to a subset of available explanatory variables. The full set of explanatory variables is not used, as this is likely to result in over-fitting. Each iteration uses a random subsample of the available data and the results consolidated to a point estimate.

The categories identified by each analyst are intended to indicate the expected interquartile range of share price changes, which is closely related to the expected share price volatility. Conversely, the ML estimates are based on 500 cross-sectional iterations. Both are designed to calibrate valuations and are likely to reflect share price volatility. Volatility is itself just one way to measure risk, and we also include measures of upside and downside risk.

To address potential issues regarding the comparability of *ARisk* and *MRisk*, we regress each on share price volatility estimated over the following quarter and predict the value of *ARisk* and *MRisk*, then *ADFor* and *MDFor* respectively. This common estimation process allows us to make comparisons based on forecasts expressed on the same scale. Interestingly, the correlation between *ADFor* and *MDFor* is 0.59, higher than that for the underlying *ARisk* and *MRisk* (0.39).

Table 4.1: Variable definitions

Name	Definition	Source
$VOLATILITY_{i,t-4}$	The standard deviation of 52 trailing weekly returns, including dividends, in local currency and annualized by multiplying by the square root of 52. Estimates below 0.1 or above 0.75 are set to missing.	Datastream
$VOLATILITY_{i,t-1}$ *$VOLATILITY_{i,t+1}$*	The standard deviation of 13 trailing/leading weekly returns, including dividends, in local currency and annualized by multiplying by the square root of 52. Estimates below 0.1 or above 0.75 are set to missing.	Datastream
ARisk	Analyst risk assessment is the difference between analysts' 'consider sell' price, i.e., the price at which a sell recommendation would be triggered, and 'consider buy' price, which is the price at which a buy recommendation would be triggered (scaled by analysts' valuation estimate).	Morningstar
MRisk	Machine learning risk assessment is the interquartile range of 500 quantitative valuations generated by Morningstar quant's machine learning algorithm scaled by the machine learning valuation estimate. Estimates above 0.75 are set to missing.	Morningstar
ADFor	Forecast volatility in the subsequent quarter based on latest analyst risk assessment (ARisk).	
MDFor	Forecast volatility in the subsequent quarter based on latest machine learning risk assessment (MRisk).	
FDFor	Forecast volatility in the subsequent quarter based on controls.	
Analyst Buy, Analyst Hold, Analyst Sell	Analysts' recommendation is a set of indicator variables where A_Buy = 1 for a buy or strong buy recommendation, A_Hold = 1 for a neutral recommendation, A_Sell if an analyst assigned a sell or strong sell recommendation, and zero otherwise.	Morningstar

ML Buy, *ML Hold,* *ML Sell*	ML recommendation is a set of indicator variables where ML Buy = 1 for a buy or strong buy recommendation, ML Hold = 1 for a neutral recommendation, ML Sell = 1 for ML sell or strong sell recommendation, and zero otherwise.	
BETA	Bayesian-adjusted beta estimate based on 5 prior years of monthly data, where Bayesian beta = 0.4 + .6 (estimated coefficient). Estimates below zero or above 2.25 set to missing.	Datastream
FIRMSIZE	Firm size measured by US$ market capitalization of equity.	Datastream
IBES_CV	Target price coefficient of variation measured by the standard deviation of target price estimates divided by the mean target price. Estimates greater than 0.4 are set to missing.	I/B/E/S
INDUSTRY	Global Industry Classification Standard (GICS) level 2 industry classification	Datastream
COUNTRY	Country of listing	Morningstar

Table 4.2: Sample by legal system

	Comparison sample				Extension sample (ML risk indicators only)			
	Common	Code	US	All	Common	Code	US	All
2012q2	111	89	412	612	561	4,527	34	5,122
2012q3	121	97	412	630	765	5,373	32	6,170
2012q4	138	86	419	643	988	5,571	31	6,590
2013q1	163	102	400	665	994	5,704	35	6,733
2013q2	224	110	393	727	664	5,819	34	6,517
2013q3	218	113	415	746	754	6,093	35	6,882
2013q4	225	109	415	749	1,032	6,205	37	7,274
2014q1	219	115	421	755	814	6,276	37	7,127
2014q2	240	116	427	783	1,211	6,840	38	8,089
2014q3	244	119	428	791	1,173	6,493	43	7,709
2014q4	253	122	430	805	1,212	5,861	44	7,117
2015q1	253	121	433	807	1,238	6,395	47	7,680
2015q2	261	123	440	824	1,264	5,428	47	6,739
2015q3	250	122	432	804	1,223	5,204	46	6,473
2015q4	258	131	437	826	1,317	5,843	42	7,202
2016q1	252	149	426	827	1,368	6,603	47	8,018
2016q2	259	156	452	867	1,385	7,569	55	9,009
2016q3	262	165	466	893	1,396	7,854	59	9,309
2016q4	275	182	473	930	1,428	7,889	60	9,377
2017q1	282	181	482	945	1,452	7,907	52	9,411
2017q2	281	179	469	929	1,436	7,294	53	8,783
2017q3	271	180	476	927	1,418	6,867	49	8,334
2017q4	264	177	473	914	1,396	6,267	51	7,714
2018q1	268	185	427	880	1,403	7,475	50	8,928
2018q2	280	186	486	952	1,536	8,253	60	9,849
2018q3	279	182	482	943	1,519	8,612	66	10,197
Total	6,151	3,597	11,426	21,174	30,947	170,222	1,184	202,353

Table 4.3: Descriptive statistics

Panel A: Sample for which both analyst and machine learning based risk assessments are available

	N	Mean	SD	Min	Max
$VOLATILITY_{i,t+1}$	21,174	0.240	0.099	0.071	0.749
$VOLATILITY_{i,t-1}$	21,174	0.238	0.099	0.067	1.179
$VOLATILITY_{i,t-4}$	21,174	0.239	0.099	0.071	0.749
ARisk	21,174	0.767	0.181	0.439	1.765
MRisk	21,174	0.142	0.078	0.038	0.747
BETA	21,174	0.986	0.363	0.002	2.248
FIRMSIZE	21,174	16.366	1.333	12.397	20.810
IBES_CV	21,174	0.109	0.051	0.000	0.400
ADFor	21,174	0.239	0.057	0.121	0.497
MDFor	21,174	0.239	0.055	0.148	0.761
FDFor	21,174	0.239	0.086	0.046	0.748

Panel B: Sample for which only machine learning based risk assessments and I/B/E/S target prices are not available

	N	Mean	SD	Min	Max
$VOLATILITY_{i,t+1}$	112,793	0.316	0.133	0.070	0.750
MRisk	112,793	0.285	0.114	0.050	0.750
MDFor	112,793	0.316	0.055	0.144	0.623
FDFor	112,793	0.316	0.096	0.086	0.831
$VOLATILITY_{i,t-1}$	112,793	0.341	0.137	0.070	1.000
$VOLATILITY_{i,t-4}$	112,793	0.318	0.136	0.070	0.750
FIRMSIZE	112,793	13.020	1.172	6.596	19.225
BETA	112,793	0.864	0.434	0.000	2.250

Panel C: Sample for which only machine learning based risk assessments and I/B/E/S target prices are available					
	N	Mean	SD	Min	Max
$VOLATILITY_{i,t+1}$	89,560	0.299	0.113	0.070	0.749
MRisk	89,560	0.244	0.110	0.050	0.750
MDFor	89,560	0.298	0.054	0.144	0.623
FDFor	89,560	0.297	0.084	0.073	0.785
$VOLATILITY_{i,t+1}$	89,560	0.313	0.109	0.071	0.995
$VOLATILITY_{i,t+1}$	89,560	0.298	0.114	0.070	0.750
FIRMSIZE	89,560	14.297	1.229	9.182	19.499
BETA	89,560	0.963	0.419	0.000	2.250
IBES_CV	89,560	0.141	0.102	0.000	3.016

Next Qtr Vol, Last Year Vol and Last Qtr Vol are the standard deviation of weekly returns, annualized, for the preceding and following year or quarter respectively. AU and QU are the analyst and machine learning risk assessments; Beta is the estimated sensitivity of share prices to market movements based on the preceding three years with Bayesian adjustment; Log(MktCap) is the log of market capital measured in US$; TP CV is the coefficient of variation of target prices recorded by I/B/E/S; and ADFor, MDFor and FDFor are the forecasts of volatility based on analysts, machine learning and financial data respectively.

Table 4.4: Correlation matrix

Panel A: Sample for which both analyst and machine learning based risk assessments are available (N = 21,174)

	VOLATILITY$_{i,t+1}$	*VOLATILITY*$_{i,t-4}$	*VOLATILITY*$_{i,t-1}$	*ARisk*	*MRisk*	*BETA*	*FIRM SIZE*	*IBES_CV*	*ADFor*	*MDFor*
VOLATILITY$_{i,t-4}$	0.674									
VOLATILITY$_{i,t-1}$	0.680	0.983								
ARisk	0.461	0.472	0.475							
MRisk	0.448	0.444	0.444	0.383						
BETA	0.252	0.268	0.274	0.285	0.214					
IBES_CV	0.489	0.538	0.539	0.440	0.402	0.243	-0.137			
ADFor	0.474	0.588	0.591	0.801	0.341	0.208	-0.191	0.385		
MDFor	0.454	0.545	0.544	0.297	0.804	0.137	-0.126	0.340	0.592	
FDFor	0.637	0.912	0.918	0.442	0.406	0.290	-0.329	0.588	0.669	0.617

Panel B: Sample for which machine learning, but not analyst, based risk assessments are available and for which I/B/E/S target prices are not available (N = 112,793 including 243 US cases not included in regressions)

	VOLATILITY$_{i,t+1}$	*MRisk*	*MDFor*	*FDFor*	*VOLATILITY*$_{i,t-4}$	*VOLATILITY*$_{i,t-1}$	*FIRMSIZE*
MRisk	0.338						
MDFor	0.346	0.862					
FDFor	0.618	0.421	0.548				
VOLATILITY$_{i,t-4}$	0.543	0.414	0.409	0.822			
VOLATILITY$_{i,t-1}$	0.605	0.381	0.434	0.920	0.750		
FIRMSIZE	-0.031	-0.286	-0.268	-0.058	0.005	-0.009	
BETA	0.137	0.172	0.175	0.245	0.362	0.242	0.044

Panel C: Sample for which machine learning, but not analyst, based risk assessments are available and for which I/B/E/S target prices are available (N = 89,560 including 941 US cases not included in regressions)

	VOLATILITY $_{i,t+1}$	*MRisk*	*MDFor*	*FDFor*	*VOLATILITY* $_{i,t-4}$	*VOLATILITY* $_{i,t-1}$	*FIRMSIZE*	*BETA*
MRisk	0.436							
MDFor	0.445	0.852						
FDFor	0.634	0.464	0.613					
VOLATILITY $_{i,t-4}$	0.578	0.500	0.486	0.807				
VOLATILITY $_{i,t-1}$	0.624	0.457	0.512	0.914	0.773			
FIRMSIZE	-0.194	-0.271	-0.245	-0.225	-0.213	-0.191		
BETA	0.132	0.201	0.158	0.173	0.276	0.196	-0.008	
IBES_CV	0.264	0.266	0.240	0.299	0.347	0.310	-0.036	0.115

Table 4.5 Sample derivation

Comparison sample: Cases with both analyst and machine learning data	Comparison sample		Extension sample	
Extension sample: Cases with machine learning but not analyst data				
Firm quarters with US$ market value plus both analyst and ML risk assessments		25,670		253,578
Missing lagged or leading volatility	1,492		12,511	
Missing beta	2,022		847	
Missing I/B/E/S target price	60		n/a	
Missing generated forecasts	346	3,920	346	13,704
		21,750		239,874
Outliers eliminated		576		37,521
		21,174		202,353
US		11,426		1,184
Code Law		3,597		170,222
Common Law (XUS)		6,151		30,947
With I/B/E/S target price available		21,174		89,560
Without I/B/E/S target price available		n/a		112,793

Table 4.6: Sample distribution by industry

	Extension sample		Comparison sample	
	N	%	N	%
Auto & Parts	6,793	3.36	482	2.29
Banks	9,249	4.57	1,105	5.24
Basic Resources	9,872	4.88	858	4.07
Chem. &	11,519	5.69	548	2.6
Construct.	12,364	6.11	609	2.89
Financials	8,392	4.15	1,174	5.57
Food & Beverage	11,217	5.54	720	3.42
Healthcare	12,714	6.28	2,047	9.71
Ind. Goods	39,323	19.43	2,867	13.61
Insurance	2,977	1.47	767	3.64
Media	4,285	2.12	480	2.28
Oil & Gas	4,535	2.24	1,521	7.22
Personal & House	11,696	5.78	1,108	5.26
Real Estate	13,482	6.66	1,111	5.27
Retail	10,202	5.04	1,613	7.65
Technology	17,131	8.47	1,539	7.3
Telecom	2,358	1.17	601	2.85
Travel & Leisure	7,922	3.91	665	3.16
Unclassified	65	0.03	102	0.60
Utilities	6,257	3.09	1,257	5.97
Total	202,353	100	21,174	100

Table 4.7: Sample distribution by country

Extension sample			Comparison sample		
Country	N	%	Country	N	%
China	44,557	22.02	United States	12,221	57.72
Japan	41,355	20.44	Australia	2,889	13.64
Taiwan	14,910	7.37	Canada	1,041	4.92
India	12,103	5.98	UK	944	4.46
Korea	8,281	4.09	France	604	2.85
UK	7,650	3.78	Japan	591	2.79
Germany	4,531	2.24	New Zealand	488	2.3
Sweden	4,413	2.18	Germany	430	2.03
France	4,150	2.05	Switzerland	338	1.6
South Africa	3,466	1.71	Netherlands	281	1.33
Indonesia	3,458	1.71	Singapore	256	1.21
Brazil	3,070	1.52	Italy	218	1.03
Switzerland	2,902	1.43	Denmark	164	0.77
Malaysia	2,802	1.38	Spain	158	0.75
Australia	2,701	1.33	Sweden	143	0.68
Singapore	2,700	1.33	Belgium	124	0.59
Thailand	2,520	1.25	Norway	63	0.3
Italy	2,171	1.07	Finland	60	0.28
Turkey	2,130	1.05	Korea	54	0.26
Poland	2,060	1.02	Taiwan	44	0.21
Mexico	1,941	0.96	Portugal	26	0.12
Norway	1,899	0.94	South Africa	25	0.12
Saudi Arabia	1,860	0.92	Mexico	12	0.06
Israel	1,816	0.9			
Chile	1,808	0.89			
Russia	1,572	0.78			
Spain	1,429	0.71			
Belgium	1,385	0.68			
Denmark	1,254	0.62			
Finland	1,238	0.61			
United States	1,184	0.59			
	189,316	93.55			
48 other countries	23,037	6.45			
Total	202,353	100.00	Total	21,174	100.00

The second Morningstar variable is a measure of dispersion within the valuation model, calculated as the interquartile range from 500 trees in the random forest. It therefore reflects uncertainty but is not explicitly a prediction of expected risk. In essence, a model is estimated which relates the predicted variable to a subset of available explanatory variables. The full set of explanatory variables is not used, as this is likely to result in over-fitting. Each iteration uses a random subsample of the available data and the results consolidated to a point estimate.

The categories identified by each analyst are intended to indicate the expected interquartile range of share price changes, which is closely related to the expected share price volatility. Conversely, the ML estimates are based on 500 cross-sectional iterations. Both are designed to calibrate valuations and are likely to reflect share price volatility. Volatility is itself just one way to measure risk, and we also include measures of upside and downside risk.

To address potential issues regarding the comparability of ARisk and MRisk, we regress each on share price volatility estimated over the following quarter and predict the value of *ARisk* and *MRisk*, then *ADFor* and *MDFor* respectively. This common estimation process allows us to make comparisons based on forecasts expressed on the same scale. Interestingly, the correlation between *ADFor* and *MDFor* is 0.59, higher than that for the underlying *ARisk* and *MRisk* (0.39).

4.5 RESULTS

4.5.1 Research Question 1: Can ML mimic analyst risk assessments?

In Table 4.8 we report models of the determinants of volatility and risk assessments. The first three columns show determinants of volatility and the raw risk assessments provided by Morningstar. Forecasts derived from our Fama-MacBeth estimation are shown on the rightmost columns. In all models, realised volatility over the preceding quarter, beta and target price dispersion are significantly positively related to the outcome variable. Company size is inversely related to risk in models of volatility and analyst risk assessments. ML risk assessments are not related to company size. Analysis of US and non-US samples (not tabulated) confirms only subtle differences in the significance of variables.

Since the factors affecting raw risk assessments and derived forecasts are consistent, we proceed with the latter. In general, we conclude that the risk assessments tend to be influenced by the same characteristics as volatility, but to different extents. There is scope to improve both sets of risk assessments; for example, the analysts might pay more attention to 12-month realised volatility, and ML might be trained on a wider set of features or explanatory variables.

We considered modelling option-implied volatility. The inclusion of VIX, a market-level option-implied volatility measure, led to no significant improvement in explanatory power. The Fama-MacBeth approach relies on cross-sectional estimation and, since implied volatility data is sparse outside the US and certain large-cap stocks in other markets, we drop implied volatility from our analysis.

We evaluate the explanatory power of the variables and the models as a whole by comparing the t-statistics and R-squared. The model of next quarter volatility has a higher explanatory variable than analyst or ML outputs, either in raw or standardized form, and this can most likely be explained by differences in the loading on realised volatility. This does not necessarily imply that risk assessments are inferior, as unmodelled elements may be related to future volatility. For example, the analyst may incorporate contextual information regarding the sector outlook. The analysis of determinants does, however, suggest that risk assessments underemphasise factors which are known to predict future volatility.

Joos et al. (2016) use panel analysis to assess the importance of accounting characteristics, specifically leverage, book to market, return on investment, negative earnings, negative equity and combinations of these variables. Our variable selection is based on this selection. In many instances, these variables are statistically significant but their contribution to the explanatory power of the models is slight, none are robustly significant across all models, and sometimes the sign switches despite being statistically significant. This lack of stability within the accounting variables may be caused by the relatively high correlation between them. This is typical for such ratios. However, the significance of non-accounting variables is unchanged by the inclusion of accounting characteristics. Previous researchers also found their results to be unaffected by the inclusion or exclusion of these accounting controls. For the sake of brevity, we exclude the accounting variables and retain realised volatility, beta, target price dispersion and company size.

4.5.2 Research Question 2: Are analyst or ML risk assessments more informative?

The workhorse for our analysis of information content is the Fama-MacBeth procedure (Fama and MacBeth 1973). The results allow an examination of the stability of the relationship and eliminate forward-looking bias. We test the influence of firm-specific financial variables but find them to be unimportant and drop them from the reported tables for the sake of brevity.

$$VOLATILITY_{i,t+1} = \alpha + \beta_1 XFor_{i,t} + \beta_2 VOLATILITY_{i,t-4} + \beta_3 VOLATILITY_{i,t-1} + \beta_4 FIRMSIZE_{i,t} + \beta_5 BETA_{i,t} + \beta_6 IBES\_CV_{i,t} + \varepsilon_{i,t}$$

To identify the relative information content of the two spread measures, we estimate the results incorporating analyst and ML spread separately and then together, with control variables included in both cases.

Table 4.8: Information content of next quarter volatility (comparison sample)

LHS is $VOLATILITY_{i,t+1}$

	1	2	3	4	5	6	7	8
ADFor	1.011***				0.754***	0.257***		0.217***
	(38.94)				(28.24)	(12.19)		(11.87)
MDFor		1.044***			0.712***		0.296***	0.258***
		(24.72)			(17.16)		(9.18)	(8.09)
FDFor			0.988***					
			(46.55)					
$VOLATILITY_{i,t-4}$				0.159***		0.150***	0.137***	0.133***
				(3.57)		(3.70)	(3.14)	(3.29)
$VOLATILITY_{i,t-1}$				0.415***		0.385***	0.393***	0.370***
				(9.22)		(8.72)	(8.08)	(7.82)
FIRMSIZE				-0.0069***		-0.0062***	-0.0066***	-0.0060***
				(-10.13)		(-8.50)	(-7.98)	(-7.12)
IBES_CV				0.322***		0.275***	0.272***	0.238***
				(24.84)		(20.26)	(16.94)	(15.80)
BETA				0.0174***		0.0123***	0.0149***	0.0108***
				(5.05)		(3.98)	(4.72)	(3.66)
Intercept	-0.00040	-0.00916	0.00868	0.164***	-0.109**	0.110***	0.106***	0.0675***
	(-0.08)	(-1.01)	(2.01)	(13.47)	(-11.16)	(7.16)	(5.98)	(3.41)
N	21,174	21,174	21,174	21,174	21,174	21,174	21,174	21,174
avg. R^2	0.244	0.215	0.528	0.558	0.330	0.570	0.574	0.583

t-statistics in parentheses, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4.8 presents results from the following Fama-MacBeth estimation, is:

$VOLATILITY_{i,t+1} = \alpha + \beta_1 XFor_{i,t} + \beta_2 VOLATILITY_{i,t-4} + \beta_3 VOLATILITY_{i,t-1} + \beta_4 FIRMSIZE_{i,t} + \beta_5 BETA_{i,t} + \beta_6 IBES\_CV_{i,t} + \varepsilon_{i,t}$
Where $XFor_{i,t}$ is $ADFor_{i,t}$ or $MDFor_{i,t}$ or $FDFor_{i,t}$

$VOLATILITY_{i,t+1}$ is volatility for the forthcoming quarter; XFor represents $ADFor_{i,t}$, $MDFor_{i,t}$ or $FDFor_{i,t}$, respectively forecasts of forthcoming vo analyst-, ML- and financial-based risk assessments at time *t* and the relationship between those variables at *t-1* and volatility at *t*; $VOLATILITY_{i,t-1}$ $_4$ are the last quarter and last year volatility; $FIRMSIZE_{i,t}$ is the log of market capitalization in US dollars; $BETA_{i,t}$ is the Bayesian-adjusted beta es 36 months using either the US market index for US firms or the FTSE world (ex-US) index for non-US firms; and $IBES\_CV_{i,t}$ is the coefficien I/B/E/S reported target price.

Table 4.9: Difference in predictive ability across legal systems (comparison sample)

LHS variable is $VOLATILITY_{i,t+1}$

	Common (XUS)	Code	Common (US)	Common (XUS)	Code	Common (US)	Common (XUS)	Code
ADFor	1.062***	0.534***	1.138***				0.756***	0.372***
	(33.05)	(10.34)	(28.44)				(17.48)	(8.68)
MDFor				1.063***	0.620***	1.281***	0.737***	0.499***
				(18.73)	(12.54)	(20.44)	(10.87)	(9.58)
Intercept	-0.0185**	0.108***	-0.0241**	-0.0198	0.0893***	-0.0587***	-0.124***	0.0255
	(-2.63)	(10.11)	(-2.75)	(-1.45)	(7.31)	(-4.81)	(-9.05)	(2.04)
N	6,151	3,597	11,426	6,151	3,597	11,426	6,151	3,597
ave. R^2	0.243	0.114	0.299	0.234	0.165	0.234	0.348	0.214

t-statistics in parentheses, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

ADFor is forecast volatility in the subsequent quarter based on latest analyst risk assessment (ARisk).

MDFor is forecast volatility in the subsequent quarter based on latest machine learning risk assessment (MRisk).

Table 4.10: Predictions of next quarter volatility (extension sample)

	Cases without I/B/E/S cover LHS variable is $VOLATILITY_{i,t+1}$			Cases with I/B/E/S cover LHS variable is $VOLATILITY_{i,t+1}$			
MDFor	0.995*** (22.82)		0.211*** (6.02)	1.106*** (19.57)		0.320*** (11.02)	0.303*** (10.91)
FDFor		0.974*** (44.82)	0.923*** (43.84)		1.003*** (38.58)	0.905*** (32.89)	0.886*** (32.85)
IBES_CV							0.0755*** (11.38)
Intercept	0.00266 (0.19)	0.0113*** (2.96)	-0.0384*** (-4.22)	-0.0321 (-1.74)	0.00383 (0.68)	-0.0641*** (-6.86)	-0.0606*** (-6.71)
N	112,793	112,793	112,793	89,560	89,560	89,560	89,560
ave. R^2	0.118	0.412	0.418	0.187	0.441	0.457	0.459

t-statistics in parentheses, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The variable definitions are the same as for the previous model. In this instance, we do not include the accounting variables. They are often individually significant, especially in the absence of prior volatility, market capitalization and beta variables, but make no substantive difference to the results.

These results show that a) analyst and ML risk assessments each contain information about future volatility, but b) both risk assessments omit relevant information. Since the control variables retain significance in the full model (columns 4, 6, 7 and 8), risk assessments are in one sense inefficient because they do not incorporate this information. Moreover, they do not fully incorporate the information included in the rival measure. Morningstar could point out that predicting volatility is not the function, or at least not the main function, of either analysts' risk categories or ML valuation variability. Even so, both analysts and ML measures are statistically significant in the presence of the other, and inclusion of either or both risk assessments leads to greater explanatory power. Both sources are useful in assessing risk.

4.5.3 Research Question 3: Do independent analysts produce unbiased forecasts of risk?

Our sample comes from independent analysts and should not display the strong biases typically found among sell-side analysts (Barber et al. 2007). Even so, there is no reason to suppose that our analysts will not be subject to the behavioural biases which typically affect decision-makers. For example, Groysberg et al. (2008) find strong optimism bias among buy-side analysts. As the analysts in our sample have to assign a recommendation from one to five to each stock, we can identify firms which are favoured as buy opportunities. If that benign attitude to the investment influences the assessment of spread, we might expect analysts to bias their spread measure down. To investigate this, we estimate the following relationship where the spread is conditioned by buy, hold and sell categories, and $ADFor_{i,t}$, $MDFor_{i,t}$ and $FDFor_{i,t}$, i.e. the derived analysts-, ML- and financial-based forecasts of forthcoming volatility based on variables at time t.

We also consider relative forecasting power across different environments. The US sample is stable and relatively homogenous and should be easier to analyse using either traditional or machine-learning methods. In additional analysis (not shown) we do indeed find that for all models the explanatory power is greater for US firms than non-US firms.

Independent analysts should not display the strong biases typically found amongst sell-side analysts (Barber et al. 2007). Even so, there is no reason to suppose that our analysts will not be subject to the general behavioral biases that typically affect decision-makers. For example, Groysberg et al. (2008) find strong optimism bias among buy-side analysts. One of the main strengths of the ML approach is the ability of the machine to extrapolate the insights of analysts to extensive samples for which analyst coverage is not available. We find that ML shows promise in challenging regulatory and economic areas.

Next, we investigate analyst bias by estimating the following relationship where the risk assessment is allowed to vary for buy, hold and sell stocks:

$$VOLATILITY_{i,t} = \alpha_0 + \beta_{1,2\ or\ 3} REC_{i,t} * (ABFor\ or\ MBFor\ or\ FBFor_{i,t}) + \varepsilon_{i,t}$$

Where *REC* is the analyst or ML recommendation, denoted as $Sell_{i,t}$, $Hold_{i,t}$ and $Buy_{i,t}$ in Table 4.10. If the analysts are unbiased, we would not expect β_1, β_2 and β_3 to be significantly different. A positive bias towards stocks categorised as buy would increase β_3, as the indicator for forthcoming risk would be relatively low and would need to be multiplied by a larger coefficient in order to forecast volatility. Here the variables represent the analyst, or ML, recommendations. Following prior research (Barber et al. 2006), we have merged strong sell with sell, and strong buy with buy, as there are few cases in the extreme categories. If analysts, or ML, are biased towards stocks they favour, we would expect the beta coefficients to be significantly different and $ß_3$ to be the highest: the low risk indicator for the stocks they favour needs to be multiplied by a larger coefficient to produce a reliable forecast of future volatility.

Table 4.10 reports our results for analyst-based and ML-based predictions of volatility conditioned by analyst and ML recommendations. In column one we report the result for analyst risk conditioned by analyst recommendation. The coefficient on *ADFor**Analyst Buy is some 10% higher than for Analyst Hold or Analyst Sell. The difference is statistically significant. Across the six columns we report, three based on analyst recommendations and three based on ML recommendations, with three different forecasts of risk (analyst, ML and financials), in only one case is the Forecast*Analyst Buy not the highest coefficient. That case is ML forecasts and ML recommendations. Our results therefore indicate analyst bias. Since the analysts in our study do not face brokerage conflicts, such as incentives to attract dealing commissions, we infer that this bias is behavioural. By contrast, ML makes unbiased risk assessments.

If the analysts are unbiased, we would not expect β_1, β_2 and β_3 to be significantly different. A positive bias towards stocks categorised as buy would increase β_3, whereas a negative attitude towards sell stocks would decrease β_1. We also run the model in the absence of fixed effects to ensure that the relationship between spread and volatility is a direct test, and the results are robust. Here the variables are as before, plus $\textit{Analyst Sell}_{i,t}$, $\textit{Analyst Hold}_{i,t}$ and $\textit{Analyst Buy}_{i,t}$ represent recommendations made by analysts or ML. For the analysts, we have merged strong sell with sell, and strong buy with buy, following Barber et al. (2007), as there are few cases in the extreme categories. This is also consistent with the ML version, which provides three categories: fairly, over- and under-valued. Buy, hold and sell are distributed 27%, 50% and 23% by the analysts and 32%, 38% and 30% by the ML system.

4.5.4 Research Question 4: How can investors best use analyst and ML risk assessments?

It is apparent that when modelled together, analyst and ML risk assessments are both statistically significant and produce higher explanatory power than either alone. A possible reason is differential effectiveness between US and non-US samples. The US sample is stable and relatively homogenous and should be easier to analyse using either traditional or ML methods. In Table 4.9 we do indeed find that explanatory power, i.e. average R-squared, tends to be greater in the US sample.

We investigate this further by dividing the non-US sample into common law and code law countries. For Australia, Canada, Hong Kong, New Zealand, South Africa and the United Kingdom, the results follow the US model and are shown in table 4.11. The predictive ability of analyst and ML risk assessments in common law countries is similar when modelled separately. When used together, *ADFor* is marginally more significant and therefore marginally more effective in predicting future volatility than *MDFor*. This would not surprise practitioners who view the operating practices of common law capital markets as having much in common. However, for the sample of code law countries, dominated by Europe plus Japan, we find that *ADFor* and *MDFor* have similar significance, and when used jointly, *MDFor* is marginally more significant than *ADFor*. We notice a deterioration in the value of analysts' risk assessments, while the significance of *MDFor* remains fairly constant. If the work of investment analysts is more difficult in code law countries, they may struggle to match the insights of analysts based in common law countries. The systematic nature of ML might be immune to such difficulties.

Table 4.11

Panel A

Analyst predictions (ADFor) conditioned by analyst recommendation (Analyst sell, Analyst hold or Analyst buy) and by ML recommendation (ML sell, ML hold or ML buy)

	VOLATILITY $_{t+1}$	VOLATILITY $_{t+1}$
Analyst sell * ADFor	0.923^{***}	
	(18.72)	
Analyst hold * ADFor	0.904^{***}	
	(19.66)	
Analyst buy * ADFor	1.004^{***}	
	(21.19)	
ML sell * ADFor		0.853^{***}
		(19.38)
ML hold * ADFor		0.874^{***}
		(18.26)
ML buy * ADFor		0.986^{***}
		(20.95)
Intercept	0.0157	0.0253^{**}
	(1.44)	(2.41)
N	21,072	20,755
avg. R^2	0.389	0.396
F-test	45.23^{***}	53.58^{***}

t-statistics in parentheses, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Panel B

ML prediction (MDFor) conditioned by analyst recommendation (Analyst sell, Analyst hold or Analyst buy) and by ML recommendation (ML sell, ML hold or ML buy)

	VOLATILITY $_{t+1}$	VOLATILITY $_{t+1}$
Analyst sell * MDFor	0.882***	
	(15.51)	
Analyst hold * MDFor	0.922***	
	(15.62)	
Analyst buy * MDFor	0.901***	
	(16.15)	
ML sell * MDFor		0.797***
		(14.65)
ML hold * MDFor		0.826***
		(14.72)
ML buy * MDFor		0.874***
		(15.96)
Intercept * MDFor	0.0233	0.0415***
	(1.73)	(3.23)
N	21,072	20,755
avg. R^2	0.357	0.378
F-test	7.13***	22.40***

t-statistics in parentheses, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Panel C

Predictions made using risk characteristics (FDFor) conditioned by analyst recommendation (Analyst sell, Analyst hold or Analyst buy) and by ML recommendation (ML sell, ML hold or ML buy)

	VOLATILITY $_{t+1}$	VOLATILITY $_{t+1}$
Analyst sell * FDFor	0.865***	
	(67.60)	
Analyst hold * FDFor	0.865***	
	(70.29)	
Analyst buy * FDFor	0.903***	
	(72.15)	
ML sell * FDFor		0.844***
		(61.46)
ML hold * FDFor		0.859***
		(64.56)
ML buy * FDFor		0.900***
		(74.66)
Intercept * FDFor	0.0302***	0.0327***
	(10.73)	(11.12)
N	21,072	20,755
avg. R^2	0.573	0.574
F-test	26.15***	50.69***

t-statistics in parentheses, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Our research provider uses ML to extend their stock coverage. The extension sample provides a further opportunity to explore the potential for ML. The ML model is trained to predict analyst valuation to price. For companies outside the Morningstar analyst coverage universe, the model is trained on the valuation of peer companies as defined (although not disclosed) by the model provider. Does predictive ability decline for stocks not covered by Morningstar analysts? By comparing the R-squared of ML in our comparison sample (Table 3.8, column 2) with the extension sample (Table 4.10, columns 1 and 4), we can see that the coefficient on MDFor remains significant. There is, however, a drop in overall explanatory power. The comparison sample R-squared drops from 0.215 (Table 3.8, column 2) to 0.12 and 0.18 in Table 4.10, column 1; the more marked decline in R-squared is evident for companies with no I/B/E/S coverage (Table 4.10, column 1) compared with those with I/B/E/S data (column 4). The extension sample shows ML to be more informative in the presence of sell-side analysts, i.e. where more extensive investor information is likely to be available.

4.6 CONCLUSION

We contrast the effectiveness of risk assessments derived from traditional financial analysis and ML for a large international sample drawn from 2012 to 2018. We find that independent analysts’ risk assessments are good predictors. Analysts’ assessments also retain an element of behavioural bias not found in the ML assessments.

Research into ML in investment research is relatively new. We believe our study is the first to conduct a head-to-head comparison of ML- and analyst-based financial analysis. Our results suggest that ML is effective, but for decisions where unstructured and unquantified information play a larger role, ML may find it more difficult to match the insights of analysts. As ML is already established, and we anticipate its use will grow rapidly, more research concerning its effectiveness would be welcome.

CHAPTER 5

An early assessment of the informational environment for equity investors since the announcement of new rules on paying for research

5.1 INTRODUCTION

This paper examines the informational environment at a time of substantial change in the marketplace for investment research. The UK investment management industry has rapidly adopted new methods to pay for analyst research in order to comply with a major new European regulation, the Second Markets in Financial Instruments Directive (MiFID II). The new rules are at least in part the product of the UK regulator's mission to protect customers, foster integrity and promote competition (FCA, 2017). Should a better-functioning marketplace for research emerge, we might expect supply and demand to be better matched, resulting in efficiency gains for investment managers and welfare gains for end investors.

Conversely, others have expressed concern that the informational environment will be depleted due to reduced quality and availability of research (Walker and Flood, 2018). This paper seeks to resolve these positions by asking whether analyst information has expanded or contracted.

Since the 1970s, it has been common for investment managers to "bundle" together execution and research costs and to pass both on to end investors. Regulatory intervention, centred around the 2003 Global Analyst Research Settlement, targeted sell-side firms and introduced rules to restrict analyst activities, disclose incentive bias and subsidize independent research providers (IRPs). It did not, however, address payment for investment research. Investment managers could continue to pay for research through the commissions paid for dealing shares to be bought or sold; these payments continued to be funded by end investors rather than the investment management firm itself.

Coinciding with these US measures, the UK government commissioned the former chairman of a leading buy-side firm to investigate the condition of the UK equity market. The government was explicitly advised to simplify the means of paying for research, i.e. to replace the opaque, relational system of dealing commissions with a neoclassical market system. The proposed outcome was that investment managers would pay for investment research. This would encourage efficiency because they would carefully consider which research to buy and would compare the costs and benefits of procuring broker research with that of doing their own research.

The UK regulatory approach had little impact until 2014, when strict new European rules on the use of dealing commissions to pay for research were proposed. Around this time, investment managers first reacted by adopting an optional, invoice-based system more extensively. By 2015, UK investment managers and research producers had come to expect new procedures for obtaining external research. These new rules, debated in regulatory consultation from 2014 to 2017, took effect at the start of 2018. Consequently, research producers and consumers are now required to agree prices for investment research in advance of use. Investment management firms can no longer access broker research without payment as this would constitute an inducement to trade according to MiFID II. To continue the established industry norm of charging end investors for research, the buy-side must also set research budgets in advance and meet the requirements of a Research Payment Account (RPA) which are far more stringent than the prior system of Commission Sharing Agreements (CSAs). Alternatively, they can treat research as a cost to their own business. A significant change came in the second half of 2017 when the UK asset management industry shifted almost in unison by choosing the latter approach, i.e. to pay for research, with a view to reducing regulatory risks, operational burdens and client confusion. Investment managers have been taking a closer look at the cost of research: media reports link these to the “exodus” in sell-side research (Walker and Flood, 2018) and a 10% fall in sell-side analyst headcount between 2012 and 2016 (Gordon, 2017), a period of buoyant stock market growth.

A newcomer to this market might expect to find significant change in research provision and subsequent impact on the informational environment. For example, it would be reasonable to expect to find fewer sell-side analysts, also less bias in the forecasts of those who remain, and possibly a narrower range of research providers. The quantity, quality and price of research might fall. Unfortunately, data to test these expectations is rather scattered. Additionally, many suppliers and consumers of research have become quite guarded in the information they wish to disclose. This trend is, in itself, informative: fund managers do not want to flag the providers they are using because of the franchise value of unique information networks.

In this paper, I examine the proportion of stocks which are covered by analysts (breadth of coverage) and the number of analysts covering each stock (depth of coverage). Narrower or shallower coverage would constitute evidence of a weaker informational environment. Next, I turn to the overall distribution of recommendations. A more even balance between buy and sell recommendations would indicate a stronger informational environment. Finally, I consider the role independent analysts constitute in supplementing stock coverage.

This examination is of practical importance because analysts have for decades played an important role in the discovery, interpretation and dissemination of information to investors (Bradshaw et al., 2017). The marketplace for research is large, important to many buy-side firms and costly for end investors, who, in aggregate, fund an estimated £1.5bn of external research per annum in the UK alone (FCA, 2014). It therefore cuts across themes of fairness and value for money in financial services (Kay, 2016). An early evaluation of the effect on the informational environment is therefore necessary.

Data vendors such as IBES, First Call, Factset and Bloomberg provide researchers with summary information from numerous research providers. This data, which is referred to as archival data in prior studies, is the principal and established method of studying analyst information. I find that the proportion of buy, hold and sell recommendations has remained almost constant since 2010. The tendency for analysts to issue buy recommendations persists at similar levels to those found bt Barber et al (2006).

Regarding the role if independent research, I find archival sources to be insufficient. Data vendors include almost no IRP predictions, archival methods cannot fully answer my questions about stock coverage. Rather than limiting the study to archival data, I use a second batch of data sources, each of which provides a window on a cross-section of data providers. Survey data shows, however, that many IRPs cover FTSE All-Share stocks. Perhaps more revealing is the fact that IRPs are not compelled to publish stock coverage.

The survey data is particularly useful for my third question which considers the range of ways that analysts inform investors. Most IRPs supply coverage only to existing clients rather than the market as a whole. Since many are specialized, and have focused client lists, limited diffusion of information is likely. Investment managers will often be able to obtain independent research on a given company, but only if they are an existing client. IRP research is less likely to leak into the marketplace. MiFID II prohibits the supply of "free" research. It seems that there are now more suppliers of company research than ever before, but the services of many analysts reach fewer investors than in a bundled system where research was made available to the market as a whole.

There seems little doubt that brokerage analysts are under pressure and that sell-side head count is falling. This is likely to result in reduced coverage but could also mean lower-quality research. If more independent analysts are covering companies, this could mitigate or even counteract shrinking coverage. Evidence of increasing numbers of independent analysts, whether or not they contribute to archives such as I/B/E/S, would count as evidence of an improvement to the informational environment. I do not propose to be able to make a causal link or even to assess whether the net effect on the informational environment is positive or negative. Nevertheless, demand from practitioners for a summary of the state of the informational environment was high, and in the absence of other independent sources, an earlier version of this paper was shared with the FCA and practitioners in early 2018.

5.2 PRIOR LITERATURE

5.2.1 The role of analysts in the information environment

Analysts play an important role in supporting the informational environment, via discovery, interpretation and dissemination of information (Bradshaw et al., 2017). Perhaps unsurprisingly, given the relative efficiency of the US stock market, recent content analysis shows that only one report in eight contains discovery and over one third of analysts never discover new information (Bradshaw et al., 2017). Interpretation seems to be more pervasive. Analysts also act as a conduit by broadcasting information even where a report lacks new discovery or insightful interpretation. Analyst activities often blend two or more of these three informational roles.

5.5.2 The effects of regulation on investment research

Researchers have only begun to examine the effects of MiFIDII. Prior studies do, however, examine analyst coverage in the aftermath of previous regulatory changes which affect investment research. Regulation of analysts remained stable from the deregulation of commissions in the US in 1975 until the early 2000s (Bradshaw et al., 2017). Analysts earned a higher profile during the 1990s and a few US analysts even became well known to the public (Beunza and Garud, 2007). In the early 2000s, analysts had to grapple with new US regulations covering corporate disclosure (Regulation Fair Disclosure), corporate governance (Sarbanes–Oxley) and sell-side analyst activities (Global Analyst Research Settlements, hereafter referred to as the Global Settlement). In an early paper on the impact of the Global Settlement, Barber et al. (2006) examined coverage and recommendation bias in First Call, a popular archive of US analyst recommendations. They found that sell-side analysts' optimism bias attenuates in response to the regulation's goal of reducing unjustifiable buy recommendations. This work was followed by a host of papers, reviewed in detail by Bradshaw et al. (2017), which examine the effects of regulatory changes, largely via analysts' predictive accuracy.

By updating the distribution of US analyst recommendations presented by Barber et al. (2006), Bradshaw et al. (2017) confirm a relative decline in sell (including strong sell) recommendations from 2002 onwards. They attribute this to the effect of the Global Settlement and related legislation which explicitly aimed to temper excessive optimism. The proportion of sell recommendations was extremely low throughout the 1993 to 2015 sample period (Bradshaw et al., 2017, p. 164) but reached its highest level in 2002 and 2003, matching a stock market trough in the wake of the 1990s technology stock bubble and accounting scandals, with a smaller peak in 2008 and 2009 reflecting analysts' views during and after the global financial crisis. Sell recommendations stabilized at around 6% to 8% of all US stocks covered from 2010 to 2015. US analysts' reticence to issue negative recommendations persists.

Barber et al. (2006) continue their study by grouping unaffiliated and independent analysts, creating investment strategies which follow the recommendations of each group, and comparing the profitability of these strategies to the distribution of ratings for each firm. This approach does not suit the present enquiry for several reasons. First, where the Global Settlement and associated legislation sought to address the bias towards buy recommendations, MiFID II is silent on practices relating to recommendations, or indeed other analyst outputs. Second, the lack of IRP data makes it impossible to make comparisons and, since unbundling affects both affiliated and unaffiliated brokers, it would not be appropriate to merge the latter with IRP analysts. Third, studies such as Brown et al. (2015), Imam and Spence (2016), Bradshaw et al. (2017) and Spence et al. (2019) provide much stronger evidence that recommendations, or indeed other analyst summary predictions, are of secondary importance to analysts' clients.

The lack of prior research on the effects of MiFIDII, or even scholarly papers attempting to consider what effects the regulation might have on analysts and their work, presents a gap. This study aims to take an early step in addressing that gap and considering the most effective approaches for further study.

Although the research market has attracted little interest from scholars over the years, several working papers have emerged which could signal a change. In the first, Fang et al. (2019) find fewer sell-side analysts covering European firms post MiFID II implementation. Using a sample of analysts covering companies listed in 31 European Economic Area countries between January 2015 and February 2019, the authors compare analyst coverage and accuracy before and after January 2018. This date is reasonable, some sell-side firms had reduced coverage well before 2018; in Unbundling Uncovered 2018 it was noted that other brokers might be expected a reassess coverage in 2019 or later. They find that over 300 companies cease to be covered by any analysts (in 90% of cases because a single analyst drops coverage). The analysts who exit tend to be less experienced, less accurate in predicting earnings more prone to making overly bullish recommendations. The find buy side analysts to be more numerous and more active in company conference calls. In short, there has been a slight trimming of sell-side analyst headcount and a shift of effort to the buy side. A second study, Lang et al (2019), uses a different sample and method, finds some evidence of falling stock coverage and also that analysts tend to cover fewer stocks in 2018 compared to 2015-2017. Neither study considers independent research providers (IRPs).

The third study in this area, which also became available in 2019, specifically addresses the effects of RPA adoption. Using the setting of Sweden, where several large asset managers unbundled their research payments in 2015, Pope et al. (2019) compare Sweden-domiciled and foreign-based analysts covering Swedish stocks. Analysis of the 2013-2016 sample reveals a fall in coverage, particularly for smaller firms and those with lower institutional ownership, but also an increase in research quality (measured primarily by the accuracy of each analyst's earnings forecasts). Limitations include specificity to the Swedish model, the choice of 2015 as a preview notwithstanding that the definition of RPA in the UK and other markets was unclear until 2016, and the absence of IRPs.

5.2.3 Studies of independent analysts

Few prior studies focus on independent research. Jacob and Rock (2008) find that investment bank analysts make more accurate earnings forecasts than their independent peers; this is likely to be a result of the superior resources afforded by the deep pockets of a Wall Street firm, including higher remuneration. Cross subsidies and certain types of knowledge sharing are typical and can be arranged without contravening regulations. Clarke et al. (2011) and Kadan et al. (2008) compare recommendations made by IRPs and sell side analysts before and after the imposition of new US analyst rules and note a shift towards less granular three-point recommendation scales (buy/hold/sell) rather than five-point scales, featuring strong buy and strong sell or equivalent. Price reaction to IRP recommendation changes became less informative. Taken together, these studies provide evidence that analysts were less likely to issue strong buys, which might have been inappropriate, but otherwise the impact on the informational environment was to provide less informative recommendations. Buslepp et al. (2014) confirm these findings over a longer time period and add that analysts at IRPs funded by the Global Settlement tend to make less accurate forecasts despite having greater financial resources. Barber et al. (2007) merge unaffiliated and independent analysts and therefore do not produce results on IRPs alone. To complement these archival studies, Allee et al. (2017) document target price accuracy of the equity research arm of Morningstar, the firm studied in Chapters 2 and 3. Like other research on IRPs, their results are confined to the US market.

5.2.4 Quantitative processes as an alternative to analysts

Analysts have been shown to provide higher-quality outputs when they are well resourced and cover fewer stocks (Clement, 1999). Since technology is more scalable, quantitative coverage is less constrained. Sell-side and independent firms have long provided quantitative recommendations to institutional clients. Even static quantitative models are demonstrably more effective than relying on experts (see, for example, Wahlen and Wieland, 2011). An example of a dynamic approach, where machine learning (ML) is applied to continuously update algorithms, can be found in Chapters 2 and 3. Systematic approaches are also likely to be less costly:

> *"Advances in technology may also decrease the value of analysts' discovery and interpretation roles. Presumably one reason for the historical reliance of the buy-side on sell-side analysts is the cost of conducting high-quality research for a large number of stocks. As costs of research decline, the buy-side may be less willing to pay for sell-side analysts' services"* (Bradshaw et al., 2017, p. 151).

In addition to cost reduction, technological advances may improve analytical capabilities. Grennan and Michaely (2018) reveal the diversity of information which is available to professional, and often individual, investors.

Fabozzi (2008) records the growth in popularity of quantitative investing in the first half of the 2000s. This trend reversed sharply in August 2007 when illiquidity emerged as a common factor which had not been priced in. Confidence in systematic processes suffered from “widespread loss of faith in quant investment methods and those who use them” (diBartolomeo, 2013, p. 7). Although some active quant processes may have faced a decline in demand, low-cost “smart beta” processes, which apply factor methodologies and generally do not require analyst research, have become a prominent part of many institutional portfolios. Brokers’ research revenues have fallen as a result.

5.3 Research Questions

I use five research questions to examine change in the informational environment, starting with an examination of the number of stocks covered by at least one analyst (breadth). With fewer analysts in post, we can expect the number of companies covered to fall.

> ***RQ1: Do research providers cover fewer companies?***
>
> ***(breadth of coverage)***

Second, I consider the number of analysts covering each stock (depth). Investors are likely to be best serviced where multiple analysts cover any given stock. Where a single analyst covers a given company, this is most likely to be the house broker who is unlikely to convey a truly critical perspective.

> ***RQ2: Do fewer analysts cover each stock?***
>
> ***(depth of coverage)***

Since investment managers are unlikely to pay for surplus, duplicate research, we can expect fewer companies to attract a very large number of analysts. Which investment manager wants to hear from the analyst ranked 28th on a given company when there are 27 other views? Some analyst retained posts because their employer wanted to be able to provide research on a very wide universe, i.e., to demonstrate breadth even if depth was questionable. My third research question is as follows.

> ***RQ3: Are fewer companies are covered by more than 20 analysts? (excessive coverage)***

Next, I examine the distribution of analyst recommendations. If the link between execution and research is removed, sell-side analysts have less incentive to make buy recommendations. We can expect to find a higher proportion of sell and strong sell recommendations.

> ***RQ4: Has optimism bias declined?***

Some investment managers use independent analysts and it would be reasonable to expect that they may have replaced outgoing sell-side analysts. My final research question addresses this.

> ***RQ5: Do independent analysts constitute a greater proportion of stock coverage***

We are most likely to observe change in coverage of UK companies because the Financial Conduct Authority (FCA) and UK market participants have shaped the new European rules. Although the changing regulations, now enshrined in the second European Markets in Financial Instruments Directive (MiFID II), affect global markets via international firms, the earliest adoption is likely to be found in the UK. We can also expect to find change in the other European Union markets. Finally, there might be less amplified changes in the US market.

I do not claim to demonstrate causality. In time, it may be possible to do so. Even so, we should be able to provide an informed assessment of the current state of the informational environment.

5.4 RESEARCH DESIGN

5.4.1 Archival analysis

Researchers have tended to use databases collecting analyst predictions. IBES, First Call and Zacks were early in requesting forecasts from brokerage firms; Bloomberg, Factest and others now have comparable offerings. Prior studies refer to analysis of this data as archival research.

I examine the number of stocks covered in each market and the intensity of coverage, i.e. the number of analysts covering each stock.

Barber et al. (2006) present graphical evidence of aggregate changes in analyst recommendations. In their case, they examine the recommendation bias before and immediately after the Global Settlement, a regulation which was explicitly aimed at preventing inappropriate buy recommendations. In this case, the regulatory change is not explicitly aimed at reducing bias. Even so, if MiFIDII reduces surplus supply, we can expect the surviving analysts to strive for accuracy. The distribution of recommendations would therefore correspond more closely to the distribution of stock returns, i.e. with approximately the same proportion of buy and sell recommendations.

MiFID II has global ramifications. Research on US firms may be purchased by MiFID II-regulated firms, and EU companies will be researched by non-EU research analysts and the research sold to non-EU investors. Although I cannot make a complete separation between the three markets, it is reasonable to expect changes to be most evident in the UK, where the regulator has been most vigilant regarding research procurement, and least evident in the US. Rather than limiting my research to a single market, I examine the US and UK separately.

The sample comprises all companies which have been present in the FTSE All-Share index since 1996. This index comprises the large cap FTSE 100, mid cap FTSE 250 and small cap firms; it therefore represents the broad market. Few companies outside these indices are covered by analysts. Repeating the analysis on all available constituents appears to make little difference to the results. I would welcome data on small company stock coverage in order to extend the analysis.

Barber et al. (2007) blend unaffiliated analysts with IRPs due to the paucity of truly independent firms. This approach identifies non-investment bank analysts and is approptriate to their post- Global Settlement setting. In contrast, MiFID II applies to all brokers regardless of investment bank affiliation. My line of enquiry requires IRPs to be studied separately from unaffiliated brokerage analysts. Although there are many hundreds of IRPs, no more than ten IRPs appear in each of I/B/E/S and Bloomberg (July 2018). It is simply not possible to justify the use of archives for the study of IRPs stock coverage across the market.

5.4.2 Survey evidence

Although I/B/E/S is known to have a long history and wide coverage, the data availability constraint noted by Barber et al. (2007) persists. Few IRPs submit to data vendors and no comparable, dedicated database exists for IRPs. Bloomberg lists over 300 brokerage firms, but only ten IRPs. Since archives provide little information about IRP analysts, I turn to other sources to obtain a more complete picture.

Several additional sources document the expansion of independent research. A survey of a leading IRP specialist (survey A), and two major surveys (questionnaires B and C) and are both described below. The surveys, one polled from IRPs and the other from their buy-side clients, provide the most representative picture I have found on independent investment research. The surveys allow us to see the research procured by early adopters of IRP research in the UK and therefore provides a window on the emergence of IRPs. I identified these datasets during fieldwork for Chapter 2 to complement archival analysis. While it would be wonderful to base the paper on a single instrument, the more specialist nature of research pricing required me to work with those best able to elicit data from informed participants; the three surveys in this paper provide the best insight I can find into the cross-section of IRP firms. A wide poll of professionals, such as that used by Brown et al. (2016) in a study investigating buy-side research, would have collected views from industry professionals who typically have rather limited knowledge of the research marketplace.

5.4.2.1 Survey A

My first additional data source is the records of the commission management department of a brokerage firm. The firm did not have a research department and instead operated as an agency brokerage to facilitate CSAs (see Chapter 2 for further details). As a result, the brokers were experts in the CSA business because this system provided their revenue in the form of commissions. The agency broker arranged CSA payments for 22 investment managers to 216 IRPs during the nine years following the introduction of CSAs in 2006. The dataset comprises annual invoice totals with descriptions and therefore reveals the actual prices paid for research services. The broker classifies IRPs into seven categories, such as fundamental, quantitative, macroeconomic and idiosyncratic. The underlying invoices are similar to the invoices issued for work undertaken by professional services firms. No questionnaire was involved; rather, this survey contains the actual accounting records and associated data such as the agency broker’s classifications and descriptions.

5.4.2.2 Survey B

My second source is a survey conducted with expert practitioners. Since all but the largest IRPs are small and specialised, at least in comparison to brokerage firms, the distribution of research is more challenging. Fund managers may not be aware of independent firms who could be best positioned to service them; IRPs may not be able to find the most appropriate buy-side specialists who might purchase their research. Integrity Research formed a business to introduce IRPs to investment managers. Founded by veterans of the IRP industry, the firm has, to my knowledge, the widest catalogue of IRP firms and also the buy-side contacts who are most cognisant of independent research. This firm is therefore ideally positioned to obtain informed questionniare responses.

I worked with the founders of Integrity Research to set the scope of the study and on the drafting of questionnaire items. Particular attention was paid to the response grids. Questionnaires set by industry and professional bodies in this field tend to be impeded by imprecise question sets and incomplete or unclear response grids. I am comfortable that the data, and that of Survey C described below, is of a superior standard.

US-based IRP specialist Integrity Research created a 17-item questionnaire on the topic of research payment. IRPs were contacted by email between September and December 2014. From 417 IRPs, 118 firms provided anonymous responses, 62% of which were US-based and 25% European IRPs. Assuming an actual population of 500 to 1000 IRPs, this implies a response rate between 12% and 28%, which is high for a survey in the investment industry. I assisted in the design of the questionnaire but had no commercial involvement with Integrity or with any IRPs.

5.4.2.3 Survey C

Investment manager practices were surveyed by RSRCHXchange, a UK-based FinTech firm focused on written research. Using a 25-item online questionnaire in April/May 2017, a sample of 562 individual responses was obtained from around 10,000 investment managers who were approached by email. A 5% to 6% sample is not uncommon in surveys of investment managers (Brown et al., 2015). As with questionnaire B, I assisted in the design of the questionnaire scope, drafting and checking. One cofounder had extensive experience in questionnaire research and employed a professional polling company. As with survey B, I am confident that the instrument design and data collection was appropriate to the rather specialist nature of this topic.

5.5 DISCUSSION AND ANALYSIS

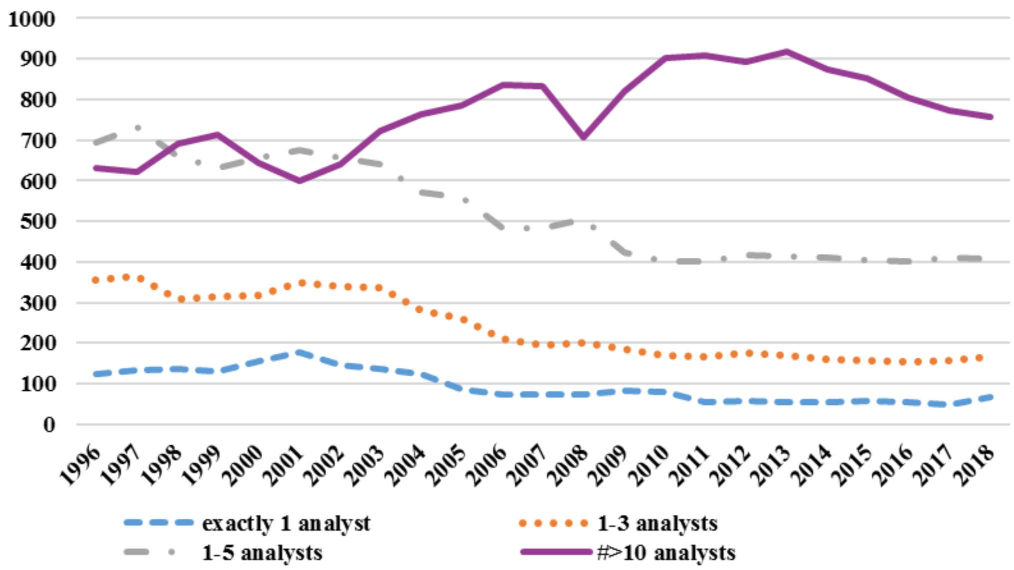
In this section, I discuss the findings for each research question based on I/B/E/S data. In the subsequent section, i.e., section 5.6, I consider additional survey evidence.

5.5.1 RQ1: Do research providers cover fewer companies?

I first address this question using archival data. Figure 5.1 Panel A shows analyst coverage of the FTSE allshare ex investment trust index, which representents large, midcap and small companies in the UK main market. almost all companies are covered by at least one analyst. In 2018, 540 companies were covered compared to an average of 542 over the past ten years, with little deviation from year to year. The level was slightly higher in the mid-2000s, peaking at 604 in 2007, before dropping to around 540 immediately after the financial crisis. After excluding investment trusts, which have corporate status but represent portfolios of other investments, almost all companies are covered by at least one analyst. (Prior to 2001, 700–800 companies were covered; the index had over 800 companies compared to around 650 post 2001.)

Figure 5.1

Panel A – US depth of coverage (S&P500)



Panel B – UK depth of coverage (FTSE Allshare ex investment trusts)

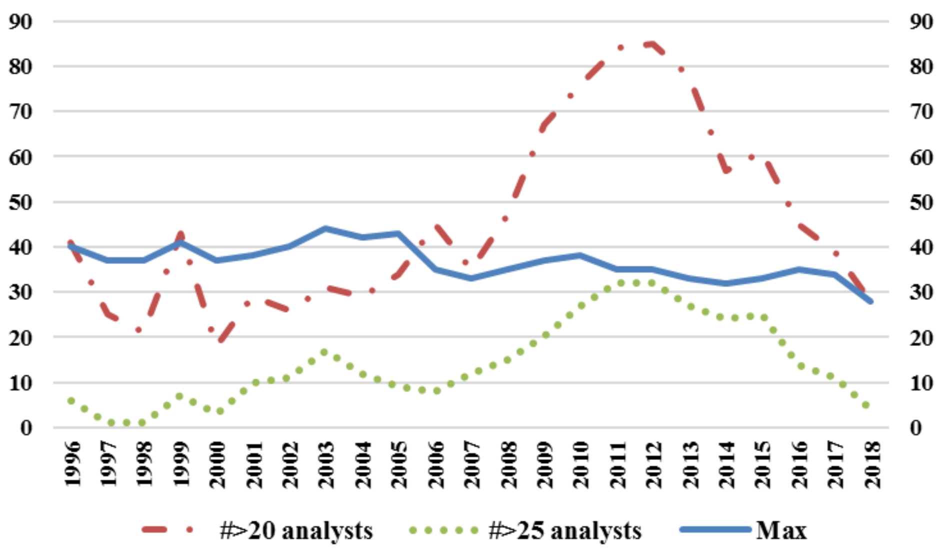


Figure 5 shows the number of companies covered by exactly one analyst, one to three analysts, one to five analysts, and more than ten analysts at each year end in the S&P 1500 (Panel A) and FTSE Allshare ex investment trusts (Panel B). The vertical axis shows the number of companies covered. In panel B (UK), the drop in number of companies covered in 2002 can be explained by FTSE's policy: fewer companies were included in the Allshare post 2002.

The number and proportion of UK companies not covered by at least one analyst appears not to have changed in the past ten years. My analysis spans the broad market index and includes small cap companies. It does not, however, include companies listed in the UK's small/micro-cap Alternative Investment Market. It may also be that sell-side firms are still taking stock of the required level of coverage. Analyst roles might be adjusted based on meetings to review 2018. Since research pricing negotiations are new, and it may take time to negotiate the price of research services, it may take several years for research provision to adjust. A complete evaluation may not be possible until the early 2020s.

Turning to the US market, the total number of stocks was highest in the tech boom, with between 1,912 and 1,952 stocks in the years 1996 to 1999 (Figure 5.1 panel B). The 1,900 level was reached again in 2004 and 2005. Coverage dropped over the last decade, dropping below 1,800 in 2013 and below 1,700 in 2015. The 2018 coverage stands at 1,560, indicating that coverage barely extends beyond the S&P1500. This marks a fall of almost 20% from peak coverage and 11% from 2015 levels. It is interesting to note that US coverage has contracted almost twice as much as UK coverage in the past decade. The downward trend is steady and does not appear to have changed pace since the announcement of new rules for research payment in 2015. Even so, the breadth of coverage remains wide in both US and UK. It may be that MiFIDII is affecting global firms but it is also possible that this trend is the result of other industry factors, such as the shift to passive investment management, pressure on buy-side fees and structural decline in brokerage commissions.

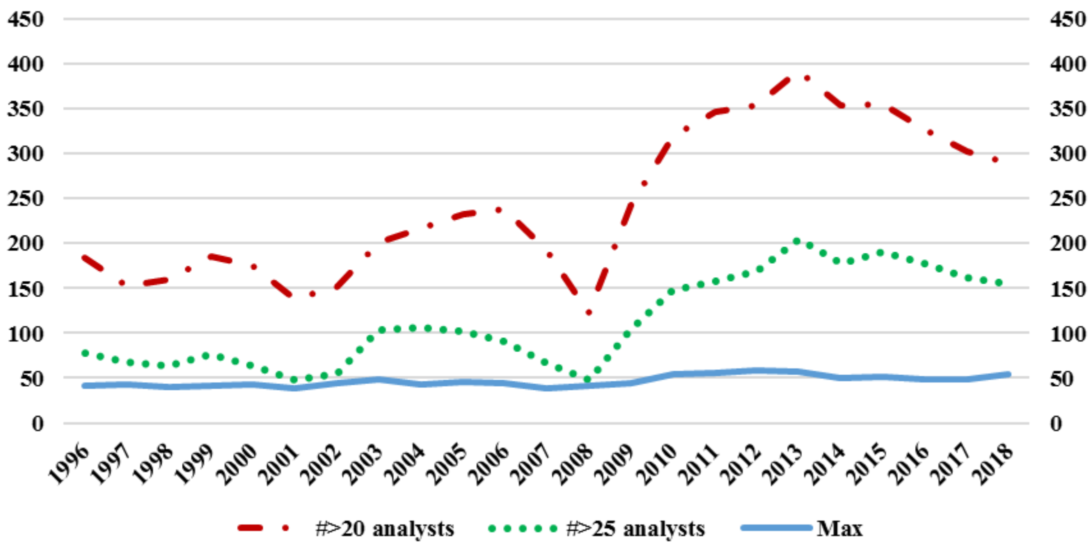
5.5.2 RQ2: Do fewer analysts cover each stock?

Investors are likely to be best served when multiple analysts cover each company. Additional analysts are likely to be less biased than the house broker and more likely to issue sell recommendations. It is therefore important to consider the depth of coverage rather than breadth alone.

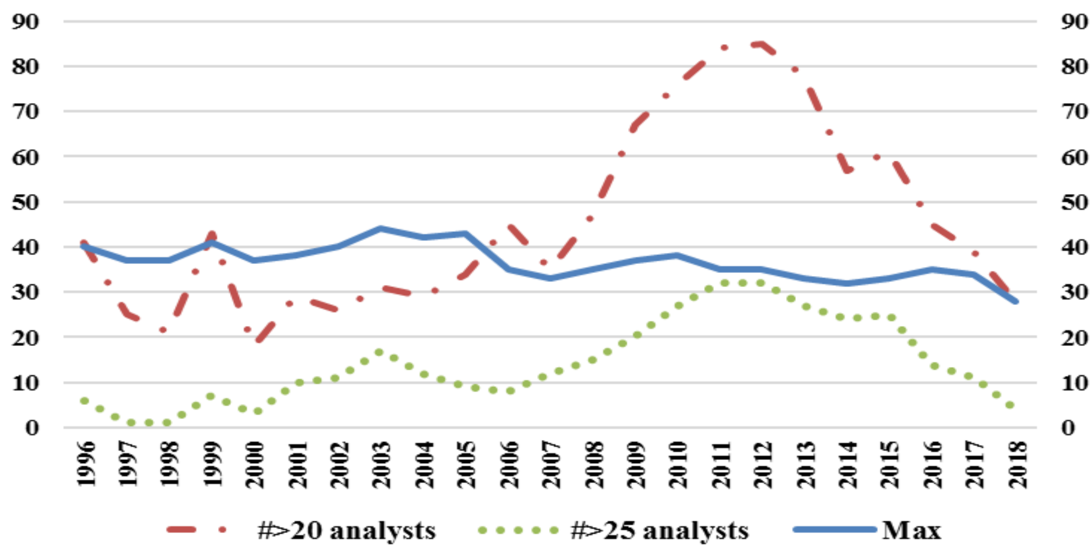
Figure 5.2 panel A shows that around 50 UK companies have been covered by a solitary analyst in each of the last ten years. This represents around 10% of the index if we disregard investment trusts. UK analyst coverage remains broad. The number of companies with sparse coverage has remained stable: almost 150 companies are covered by one, two or three analysts and approximately 200 are covered by one to five analysts. Deep coverage, which I define as the number of companies covered by more than ten analysts, declined steadily from around 250 companies in 2011 to 200 in 2018.

Figure 5.2

Panel A US



Panel B UK



Panels C and D show the number of companies covered by many (more than 20 and more than 25) analysts at year end. Max represents the maximum number of analysts covering any single company.

Turning to the US market, Figure 5.2 Panel B shows that the depth of stock coverage has been stable over the past decade. More companies received deep coverage in the aftermath of the Global Settlement. Roughly 600 companies were covered in 2001; this rose to over 900 in the early 2010s. In the past five years, we can see a trend towards shallower coverage towards 2018, when 756 companies were covered by more than ten analysts.

The combined effects of the Global Settlement (2003) and the early 2000s recession and stock market trough had a significant effect on financial services headcount in the UK. There is some evidence that coverage deepened during the 2000s: fewer companies were sparsely covered; more were deeply covered. The 2010s exhibit a stable level of deep coverage. We cannot conclude that fewer analysts cover each stock both in the UK and US.

Research providers may choose other ways to try to provide research more efficiently. Analysts may be required to cover more stocks and could dilute quality. The number of stocks covered by each analyst is, however, of limited use as a quality measure. The adoption of technologies such as ML (Chapters 2 and 3; Grennan and Michaely, 2018) could allow analysts to expand coverage and improve quality. Like Bradshaw et al. (2017), I expect this area to be an important field in research on analysts in the coming decade.

5.5.3 RQ3: Are fewer companies excessively covered?

A more significant change is evident when examining those companies which are covered by more than 20 analysts. It seems unlikely that the informational environment would be improved by additional analysts above this threshold. Since few investment managers will be prepared to pay for less valuable research, we can expect fewer highly ranked analysts to cease coverage. Around 30 UK companies are covered by 20 or more analysts.

In 2013, nearly 400 US companies were covered by more than 20 analysts and 200 were covered by more than 25 analysts. As at 2018, these figures have dropped to 150 and 300 respectively. Only five companies were covered by more than 25 analysts in 2018 compared to 25 in 2015 at the announcement of new payment rules.

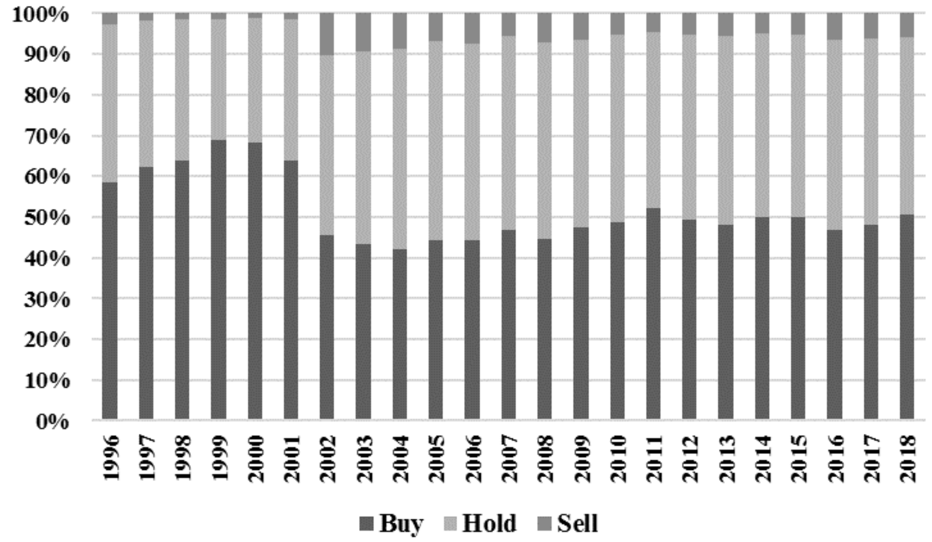
Evidence from archival sources therefore confirms that brokerage firms have trimmed some of the redundant investment research. We know that some brokers have ceased covering the UK market, and others have chosen to focus on particular sectors. Brokers may also have removed analysts who produced less revenue in the form of payments from investment management firms. It is evident that fewer companies are excessively covered. In both the UK and US, it seems clear that the termination of research coverage by any single research provider has little effect on the overall informational environment.

5.5.4 RQ4: Is optimism bias lower?

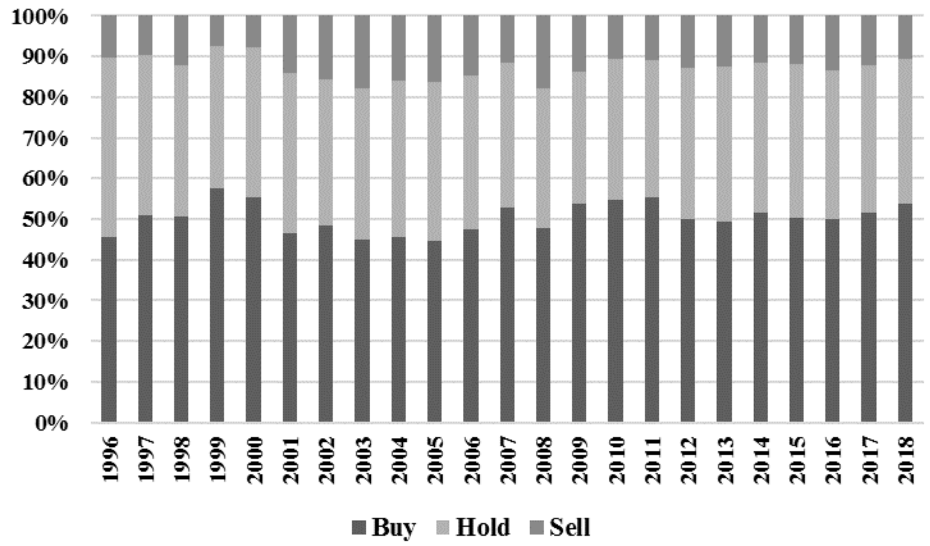
We follow the approach of Bradshaw et al. (2017), who in turn update the findings of Barber et al. (2006, 2007), by examining coverage and recommendation bias in archives of analyst predictions. Bradshaw et al. (2017) report on US analysts' recommendation categories ranging from strong buy (1) through neutral (3) to strong sell (5). From 1993 to 2000, this average was always above 3.5 and, in 2000, reflecting the exuberant technology stock boom, reached a high of 4. At its lowest, the bias disappeared (reaching 3.0, i.e. neutral) in only two years: the 2002 bear market and 2008 global financial crisis. The average recommendation was steady at around 3.5 from 2010 to 2015. The sample ends in 2015 and consequently no inference can be made regarding the introduction of MiFID II.

Figure 5.2

Panel A – US recommendation distribution



Panel B - UK recommendation distribution



In my analysis, the average recommendation across all UK companies is 2.14 (as at October 2018), very close to the ten-year average of 2.20 within a range of 2.12 to 2.30. The US sample is closely in line with Bradshaw et al. (2017) and I find no shift in the tilt towards optimism.

The relatively low percentage of buys (and high percentage of sells) in 2009 is likely to reflect the financial crisis. The proportion of buy, hold and sell recommendations has remained almost constant since 2010.

Optimism bias persists. In UK, US and European markets, the bias towards buy recommendations has remained at similar levels since the mid-2000s. I note, however, that few independent analysts supply their data to archives. It is possible that recommendations which are shared only with investment management clients follow a more balanced distribution.

The data availability constraint noted by Barber et al. (2007) remains: few IRPs submit to these databases. For example, Bloomberg lists over 300 brokerage firms but only ten IRPs. Although no comparable database for IRPs exists, a major survey indicates that over 400 IRPs exist. I combine secondary analysis of this survey data to reveal a clearer picture of the information provided by IRPs. I examine these in the next section.

5.5.5 RQ5: Do independent analysts constitute a greater proportion of stock coverage?

Since only around ten IRPs supply data vendors, we cannot use archives to answer this question. Unlike brokers, IRPs are not required to record or publish stock coverage. Bloomberg reveals the identity of all contributing firms. In July 2017, ten IRPs contributed alongside 322 brokerage firms. A search of I/B/E/S coverage for major US and European indices revealed a comparable number of IRP contributors. Taken at face value, this would indicate that IRPs are just as rare as Barber et al. (2007) found in the wake of the 2003 Global Settlement. Despite this, surveys A and B in this chapter show the number of IRPs seems to have expanded at least since the mid 2000s.

Investment managers use estimate archives such as I/B/E/S and Bloomberg to check the consensus view and the stance taken by individual analysts, but will almost certainly access reports and seek analyst interactions. A 2017 survey (questionnaire B) shows that written reports are the most frequently used and highly valued component of analysts' work. These reports are usually sent to investment managers via private email or accessed via password-protected websites hosted by individual brokers or aggregators such as Thomson Reuters or Factset. Vendor data on recommendations tends to be for reference use (e.g. to play "devil's advocate" or to discover "what the street is thinking") or by quantitative funds who use them as an input to their process.

It is clear that very few IRPs contribute their recommendations to vendors such as Bloomberg or I/B/E/S. There are some possible reasons for the sparsity of IRP recommendations in vendor databases. First, many IRPs do not produce stock recommendations: fewer than half of the IRPs in survey A make stock recommendations. Instead, these firms are specialists, for example conducting analysis on an economy, an industry, a political event or technological innovation. Second, most IRPs who do make stock recommendations choose not to supply data vendors and as a result protect their intellectual property from quickly entering the informational environment. Third, most IRP research is, to some extent, exclusive: recommendations and other services are reserved only for their paying customers, thus protecting the IRP's intellectual property and the investment manager's information franchise.

Investment managers use estimate archives such as I/B/E/S and Bloomberg to check the consensus view and the stance taken by individual analysts, but will almost certainly access reports and seek analyst interactions. A 2017 survey (questionnaire B) shows that written reports are the most frequently used and highly valued component of analysts' work. These reports are usually sent to investment managers via private email or accessed via password-protected websites hosted by individual brokers or aggregators such as Thomson Reuters or Factset. Vendor data on recommendations tends to be for reference use (e.g. to play "devil's advocate" or to discover "what the street is thinking") or by quantitative funds who use them as an input to their process.

It is clear that very few IRPs contribute their recommendations to vendors such as Bloomberg or I/B/E/S. There are some possible reasons for the sparsity of IRP recommendations in vendor databases. First, many IRPs do not produce stock recommendations: fewer than half of the IRPs in case study A make stock recommendations. Instead, these firms are specialists, for example conducting analysis on an economy, an industry, a political event or technological innovation. Second, most IRPs who do make stock recommendations choose not to supply data vendors and as a result protect their intellectual property from quickly entering the informational environment. Third, most IRP research is, to some extent, exclusive: recommendations and other services are reserved only for their paying customers, thus protecting the

IRP’s intellectual property and the investment manager’s information franchise.

5.6 SURVEY EVIDENCE

Since the Global Settlement, US brokerage analysts have been required to disclose coverage, and this practice has been mimicked globally. IRPs do not face such a requirement and Chapter 1 revealed that it is very common for the recommendations to be privately available to clients, or indeed that no stock-level recommendations exist at all.

A subset of IRPs mimic the structure of the investment research department at a brokerage firm. Some employ scores of analysts, divided by sectoral specializations; they supply investment managers with reports, predictions, calls and meetings. In the past decade, some IRP analysts have ranked among the leading analysts in investment manager surveys (e.g. Extel).

Standardization does not, however, appear to be the norm. A 2015 survey identified 417 IRPs (questionnaire B) and noted that hundreds more are likely to exist. Examining ten Bloomberg contributors is therefore unlikely to be representative of the industry. Instead, a secondary analysis of survey data shows extensive variety in the types of independent research used by investment managers.

The stock exchange rules associated with the Global Settlement require brokers to publish a distribution of recommendations. As a result, each broker's research universe, including initiation and cessation of coverage, is carefully disclosed to investment management clients. IRPs do not have the potential conflicts of interest inherent in brokerage firms, they operate without the full constraints imposed by the Global Settlement, and they are therefore free from such restrictions. This freedom allows them to provide different types of coverage. IRPs can set their universe and allocate analysts accordingly. Some augment coverage by producing research on other companies on request; others refuse to define a fixed coverage list at all: IRPs need only inform their existing clients. Merely giving away the name of a company may signal potential interest and might give non-client investment managers a new "idea". Quantitative IRPs provide yet another form of coverage: these firms generate lists or portfolios of trade ideas, valuations, sentiment measures and risk characteristics. In summary, IRPs can provide flexible, exclusive and informally defined coverage.

Other segments of the IRP industry do not attempt to provide company-specific recommendations at all. Most “macro” IRPs do not include valuations or recommendations for individual stocks. Some of the longest-established IRPs are macro, and this subsector is one of the largest, in number and aggregate revenue, in the UK market. The largest category is, in fact, labelled “idiosyncratic” by the agency broker. Other examples from the IRP invoice dataset (survey A) include political commentary, customer surveys and industry expertise. It appears that many IRPs choose strategies to complement, rather than replace, broker research. There is considerable redundancy in broker research and so it seems clear that investment managers buy analysis from IRPs; the informational environment is enhanced but the channel of transmission is not via stock coverage. Even so, IRPs provide a mix of competition and complementarity.

Bradshaw (2009) provides a five-stage illustration of a sell-side analyst’s information process. In the first three stages, the information is collected, processed and used to estimate earnings or cash flows. The final steps are valuation and recommendation. This depiction seems uncontroversial. IRPs who do not provide forecasts must, therefore, limit their activities to some combination of the first two activities, i.e., information collection and processing. This finding also aligns with investment managers’ demands for business analysis rather than recommendations (Beunza and Garud, 2007;

Imam and Spence, 2016). The buy-side analyst uses this information, then makes her forecasts, valuation and recommendation.

Gleason and Lee (2003) find a considerable lag in price adjustment following publication of research by analysts who are accurate but less well known; many IRPs are likely to fit this description.

Some IRP research is exclusive to the client, i.e. it is written for the use of a single investment management firm and is not available to competing firms. A large dataset of IRP invoices (survey A) reveals that around half of IRPs who perform fundamental research do so primarily on an exclusive basis, with some firms specializing in this type of work. Research of this type is less likely to duplicate the reports of sell-side analysts and is likely to be distilled slowly into the informational environment.

Some market participants have expressed concern that new research payment rules would reduce the availability of information to investors. In the UK, the birthplace of research unbundling, the number of stocks covered by analysts remains steady although with less excessive duplication. IRPs rarely supply data vendors with their recommendations and are not required to disclose recommendation distributions. As a result, they offer more flexibility in their ability to focus resources. There are now many more IRPs than brokers, and this may improve the informational environment for investment managers who pay for access.

5.7 LIMITATIONS

Archival data has numerous limitations, not least rather sparse IRP data. My approach complements archives and makes this limitation less serious. Even so, the following limitations should be noted.

We use archives to assess coverage but make no evaluation of the quality of the research provided. Assessing quality is challenging. Forecast accuracy has been used in prior studies but not without challenges (Beunza and Garud, 2007; Imam and Spence, 2016); my own data (Chapter 1) shows that accuracy has barely been mentioned in industry discussions on research pricing. Rankings such as Extel and Institutional Investor indicate the relative popularity of each analyst but identifying the top three or five analysts in each sector would add little to the study. Experience and resources offer some potential (Brown et al., (2015). Practitioners have voiced concern regarding a trend of veteran sell-side analysts being replaced by less experienced subordinates: this effect, sometimes termed “juniorisation”, was a recurring theme at the 2018 Substantive Research conference. Quality remains difficult to measure.

Data vendors do not typically pay research providers to do this and the quality and availability of such data is shrinking. UBS, a broker which had one of the widest and deepest research capabilities in the past three decades, is one example of a firm which ceased supplying data vendors; this change was effected in the run-up to MiFID II.

We have studied the coverage of large and midcap companies in two major markets. Investors and regulators are also concerned about the provision of information on smaller firms. US and UK stock markets may be losing prominence compared to emerging markets such as China, which are now well established and companies listed there will be attracting a new generation of analysts. Geographic classifications also add complexity. MiFID II affects the provision of research by and to firms with customers in European jurisdictions. UK investment managers and research providers often have global coverage. Conversely, some UK companies will be covered by analysts who are not based in the UK. This mismatch exists more generally, i.e. US and Europe, and for global markets. Research regulation therefore affects the information environment.

Finally, I note that regulatory change need not be exogenous. The present case, MiFID II, affects some firms but not others both in EU and non-EU informational environments. Member state regulators are not uniform in their application and enforcement. Requirements were announced in a series of communications between 2014 and 2017. In short, studies attempting to use research payment changes as an exogenous shock face greater methodological challenges than studies of the Global Settlement. I do not make claims regarding the causal nature of any change observed.

5.8 CONCLUSIONS

Analysts have for decades played an important role in the informational environment. Their world has undergone dramatic change in the past decade, most recently due to changes in the procedures used to pay for investment research, primarily due to MiFID II. A structural decline in the profitability of equity brokerage, exacerbated by regulation on research payment, means that many research providers are expected to reduce analyst headcount and produce research on a more limited selection of companies. Market participants have voiced concerns that analysts now cover fewer stocks. I present early empirical evidence on stock coverage and consider the role of IRPs.

I examine change in stock coverage since the announcement of new payment rules in the UK. Archival data shows no economically meaningful drop in the number of companies covered. I see no increase in the proportion of relatively large companies covered by only one analyst, or a small number of analysts. It appears that some of the most intensively covered stocks are now followed by fewer but still more than 20 analysts. Some surplus coverage has therefore been trimmed and it seems unlikely that this would reduce the available information set on large companies.

My own contribution is more basic than either of these papers yet is timely and has been shared with regulators in early 2019. The analysis presents early descriptive analysis on stock coverage in the US, UK and Europe. I examine change in stock coverage since the announcement of new payment rules in the UK. I/B/E/S data shows no economically meaningful drop in the number of companies covered. I see no increase in the proportion of relatively large companies covered by only one analyst, or a small number of analysts. It appears that some of the most intensively covered stocks are now followed by fewer, but still more than 20 analysts. Some surplus coverage has therefore been trimmed and it seems unlikely that this would reduce the available information set on large companies. The data reveals modest change in the number of stocks covered by analysts and no reduction in the bias towards buy recommendations. Since the number of listed companies is falling in some markets (e.g. UK) and increasing in others, I have also recently checked that I/B/E/S coverage in China and India is increasing. Stock coverage continues to follow institutional investor demand.

Archival analysis reveals modest change in the number of stocks covered by analysts, insubstantial change in breadth and depth of coverage, and no reduction in the bias towards buy recommendations.

In this chapter I investigate the effect of IRPs on stock coverage. Since archival sources contain very few IRP predictions, I use secondary analysis of three surveys to examine IRP stock coverage. Equity IRPs tend to be sector specialists. IRPs are not required to disclose coverage, and often choose not to do so. This strategy protects the value of their ideas. It is also clear that many IRPs draw softer boundaries around their coverage universe; they can initiate or cease coverage without the formal communique required by brokerage firms.

Since their definition of stock coverage is more flexible, the means by which IRPs contribute to the informational environment constitutes the third question in this study. IRPs may or may not provide stock recommendations and many focus instead on business or industry research which investment managers use to make their own valuation and decision. Indeed, the largest category of IRPs provides idiosyncratic services. The survey data also shows that independent investment research reveals different categories and that novel types of coverage exist.

CHAPTER 6

Conclusion

This thesis investigates regulatory and technological changes in investment research. As a starting point, the first empirical study identifies and explains the mechanisms used to pay for investment research over the past four decades. The reciprocal arrangements embedded in the research marketplace are structured as a gift exchange. Buyers and sellers of research have typically favoured this opaque system and resisted change.

Using data collected from specialist industry events and documents, chapter 2 of this thesis identifies five market mechanisms used to pay for investment research in the past four decades. The characteristics of each mechanism are classified according to their orientation towards neoclassical market or reciprocal exchange. The latter type of system has prevailed. Research has typically been provided without explicit pricing in order to attract a counter-gift (brokerage commission) which is proportional to trade size and therefore grows proportionally with assets under management. MiFID II banned such practice in Europe, imposing competition in its place and thus performing the marketplace in the image of an economic model.

The change has important ramifications. Investment managers must now evaluate the cost of research alongside the other costs incurred to run their business, allowing managers to compare external and internal research costs such as analyst compensation. Greater scrutiny can be expected than was typical in many reciprocal systems. Research which is not perceived to add value will be removed from future research budgets. In many cases, the overall budget for research will be lower than it was prior to 2018. Research providers must negotiate the price of research in monetary terms rather than percentage points and reach agreement in advance of use. For firms with little prior use of CSAs or P&L-paid research, this change will be radical. Few brokers now attempt to provide full coverage of listed stocks in all sectors or countries; instead, most specialize. End investors have greater transparency regarding the costs of investing. Since investment managers have a regulatory requirement to avoid paying for underutilized research, we can expect less duplication. Competition should result in greater value for money. It may be harder for investors to find and evaluate research but entrepreneurs have set up stall to address these challenges.

My empirical results indicate that the change has not affected firms in a uniform fashion; it is therefore difficult to identify an exogenous change. Regulatory change was negotiated over several decades. MiFID II is enacted somewhat differentially in each European Union member state and has also induced US regulatory change. A sharper focus on the cost of research may improve quality but also reduce quantity.

Cost pressures faced by research providers and investment managers are one motivation for automation, including the adoption of machine learning (ML). The second study – chapter 3 – examines the ability of analysts and ML to predict returns. This task is substantially harder as it requires the ability to beat markets which are thought to be reasonably efficient. Analysts' valuations show some ability to predict next quarter returns. ML valuations are contrarian, most likely because the algorithm is sensitive to short-term price movements. When used together, analyst and ML valuations show significant ability to predict returns, but since this strategy requires the ML signal to be inverted, this seems contrived and infeasible in practice. ML appears to be too sensitive to price changes, which drive the target variable which is defined as analyst valuation divided by price.

The third study - chapter 4 - presents a comparison of the relative effectiveness of risk assessments derived from traditional financial analysis and ML for a large international sample. Prior studies show that analysts are effective in predicting fundamental risk, i.e. the distribution of possible valuations around a central estimate. My analysis confirms that risk assessments provide incremental information about fundamental risk. I also contribute to the literature on the second moment of analysts' target price estimates by showing that independent analysts retain an element of behavioural bias. This bias is not found in the ML assessments.

These results have several implications for investors. The first is that ML may be able to perform tasks which have previously required human expertise. In this case we examine the calibration of target prices but it may also apply more generally to tasks where analysts demonstrate successful prediction. The second is that, because analyst and ML risk assessments are not perfectly correlated, investors will be better informed by using a combination of analyst- and ML-derived risk assessments. The third implication is that analysts may be able to monitor and recalibrate their own risk assessments, particularly where associated with their own buy recommendations, in order to reduce self-attribution bias.

My results have two implications for the work of financial analysts. First, combining the assessments from the ML process with that produced by analysts clearly improves the information content of the analysts' work. Recommendations made by our sample analysts are computed using the risk assessment, and so analysts may benefit from combining the insights from traditional investment analysis with that from ML. Second, ML learning appears to match the effectiveness of traditional analysis in producing informative risk assessments, and its cost-effectiveness leads to substantial increases in coverage and much faster updating. Particularly in the market for investment research, where regulatory changes have put considerable pressures on cost structures, cost-effective ML techniques may become widely adopted.

Research into ML in investment research is relatively new. Chapters 3 and 4 provides new evidence on the use of ML to mimic analysts. This chapter is, to my knowledge, the first to conduct a head-to-head comparison of ML- and analyst-based financial analysis. Our results suggest that ML is effective but for decisions where unstructured and unquantified information play a larger role, ML may find it more difficult to match the insights of analysts.

The third and fourth chapters also contribute to the literature on technological change in investment analysis. Much of our existing knowledge about investment analysts and their work comes from studies of sell-side analysts in the US. Chapters 3 and 4 present evidence on independent analysts in a global setting. For practitioners, the results suit global equity strategies rather than limiting the study to US equities. The key message from this study is that ML cannot be expected to succeed where expert predictions have little or no predictive ability.

The fourth study – chapter 5 – takes stock of the informational environment. The financial media has reported a decline in stock coverage. My results confirm a steady but gradual decline in the breadth and depth of stock coverage in the past decade. But there is no evidence of a sharp decline since the announcement of new regulations in 2015 or the introduction of MiFID II in 2018. Archival evidence indicates that stock coverage remains wide and deep. There has been a modest contraction in the coverage of UK stocks in the current decade compared to a slightly greater contraction in the US. Although few independent analysts contribute to archives, case study evidence shows a healthy independent sector to complement brokerage analyst research.

Perhaps the most important implication of this simple empirical analysis is that archives of analyst data are becoming less representative of the forecasts made by experts. There are several reasons for this: first, a small minority of independent experts submit their forecasts; second, some top-tier brokerage firms no longer submit analysts' forecasts; third, there are now more crowdsourced estimates and other FinTech solutions (see, for example, Grennan and Michaely, 2018); and, finally, buy-side firms increasingly develop their own capabilities. Even the most established vendors are affected by these trends. I also present evidence that there has been relatively little change to date in the availability of information to investors in two major markets. Archival sources show that the availability of sell-side analyst research has slowly contracted but no sudden drop in coverage was evident in 2018.

I make no evaluation of the quality of the research provided. Quality is hard to define and is subjective. Proxies such as forecast accuracy have become increasingly contested (Imam and Spence, 2016), and rankings may be incomplete. In chapters 3 and 4, analysts in our single-firm sample have extensive experience and resources, and this in turn provides some indication of quality. In chapter 5, analysis on US and UK broad market indices and therefore provides no evidence on small companies or those listed in other markets.

Finally, I note that regulatory change need not be exogenous. The present case, MiFID II, affects some firms but not others both in EU and non-EU informational environments. Member state regulators are not uniform in the application and enforcement. Requirements were announced in a series of communications between 2014 and 2017. In short, studies attempting to use research payment changes as an exogenous shock face greater methodological challenges than studies of the Global Settlement. I do not make claims regarding the causal nature of any change observed.

Archival data has numerous limitations. IRPs often choose not to supply archives and take a less formal approach to stock coverage. Novel types of research exist which may not even include earnings forecasts, stock recommendations or target prices, the three summary outputs upon which most scholarship is based. Archives therefore capture a diminishing share of analyst research. This thesis contributes to the rather sparse literature on independent investment analysts.

In chapter 5 I focus on broad market coverage and include the UK and US listed companies which most investment managers could hold. The analysis excludes very small companies which might be held in specialist funds. To my knowledge, there is no archive of stock coverage for very small companies, such as those on the UK Alternative Investment Market. The study could, however, be extended by examining markets where stock coverage is expanding, for example in emerging markets. Finally, recent testimony from practitioners highlights that the annual review of research costs might lead to a fall in overall spending in 2019 or even later.

Taken together, this thesis contributes to our existing knowledge of investment analysts. Specifically, it adds to the existing literature on the economics of the marketplace for analysts' work, the expanding category of independent analysts and the use of ML to mimic analysts' outputs.

REFERENCES

Allee, K. D., Erickson, D., Esplin, A., & Larocque, S. (2017). "Independent Analysts' Estimates of Firm Value". Working paper.

Akerlof, G. A. (1982). "Labor contracts as partial gift exchange." The quarterly journal of economics **97**(4): 543-569.

Arjaliès, D.-L., P. Grant, I. Hardie, D. MacKenzie and E. Svetlova (2017). Chains of finance: How investment management is shaped, Oxford University Press.

Asquith, P., M. B. Mikhail and A. S. Au (2005). "Information content of equity analyst reports." Journal of financial economics **75**(2): 245-282.

Ball, R. T. and E. Ghysels (2017). "Automated earnings forecasts: beat analysts or combine and conquer?" Management science.

Baltussen, G., B. van der Grient, W. de Groot, E. Hennink and W. Zhou (2012). "Exploiting option information in the equity market." Financial analysts journal **68**(4): 56-72.

Barber, B., R. Lehavy, M. McNichols and B. Trueman (2001). "Can investors profit from the prophets? Security analyst recommendations and stock returns." The Journal of finance **56**(2): 531-563.

Barber, B. M., R. Lehavy, M. McNichols and B. Trueman (2006). "Buys, holds, and sells: The distribution of investment banks' stock ratings and the implications for the profitability of analysts' recommendations." Journal of accounting and economics **41**(1-2): 87-117.

Barber, B. M., R. Lehavy and B. Trueman (2007). "Comparing the stock recommendation performance of investment banks and independent research firms." Journal of financial economics **85**(2): 490-517.

Barberis, N. and R. Thaler (2003). "A survey of behavioral finance." Handbook of the economics of finance **1**: 1053-1128.

Barker, R., J. Hendry, J. Roberts and P. Sanderson (2012). "Can company-fund manager meetings convey informational benefits? Exploring the rationalisation of equity investment decision making by UK fund managers." Accounting, organizations and society **37**(4): 207-222.

Barker, R. G. (1998). "The market for information—evidence from finance directors, analysts and fund managers." Accounting and business research **29**(1): 3-20.

Bell, D. (1991). "Modes of exchange: Gift and commodity." The Journal of socio-economics **20**(2): 155-167.

Beunza, D. and R. Garud (2007). "Calculators, lemmings or frame-makers? The intermediary role of securities analysts." The sociological review **55**(2_suppl): 13-39.

Bilinski, P., D. Lyssimachou and M. Walker (2012). "Target price accuracy: International evidence." The accounting review **88**(3): 825-851.

Block, S. B. (1999). "A study of financial analysts: practice and theory." Financial analysts journal **55**(4): 86-95.

Bonini, S., L. Zanetti, R. Bianchini and A. Salvi (2010). "Target price accuracy in equity research." Journal of business finance and accounting **37**(9-10): 1177-1217.

Bradshaw, M., Y. Ertimur and P. O'Brien (2017). "Financial analysts and their contribution to well-functioning capital markets." Foundations and trends in accounting **11**(3): 119-191.

Bradshaw, M., A. Huang and H. Tan (2012). "Analyst target price optimism around the world."

Bradshaw, M. T. (2009). "Analyst information processing, financial regulation, and academic research." The accounting review **84**(4): 1073-1083.

Bradshaw, M. T. (2011) "Analysts' Forecasts: What Do We Know after Decades of Work? " Working paper.

Bradshaw, M. T., L. D. Brown and K. Huang (2013). "Do sell-side analysts exhibit differential target price forecasting ability?" Review of accounting studies **18**(4): 930-955.

Brav, A. and R. Lehavy (2003). "An empirical analysis of analysts' target prices: Short-term informativeness and long-term dynamics." The journal of finance **58**(5): 1933-1967.

Breiman, L. (2001). "Random forests." Machine learning **45**(1): 5-32.

Brown, L. D., A. C. Call, M. B. Clement and N. Y. Sharp (2015). "Inside the 'black box' of sell-side financial analysts." Journal of accounting research **53**(1): 1-47.

Brown, L. D., A. C. Call, M. B. Clement and N. Y. Sharp (2016). "The activities of buy-side analysts and the determinants of their stock recommendations." Journal of accounting and economics **62**(1): 139-156.

Buslepp, W. L., R. Casey and G. R. Huston (2014). "Did they get what they paid for? The Global analyst research settlement and analyst research quality." The global analyst research settlement and analyst research quality (April 16, 2014).

Bouwman, M. J., Frishkoff, P. A., & Frishkoff, P. (1987). How do financial analysts make decisions? A process model of the investment screening decision. *Accounting, Organizations and Society*, *12*(1), 1-29.

Callon, M. (1998). "The laws of the markets." Sociological review monograph.

Carhart, M. M. (1997). "On persistence in mutual fund performance". Journal of finance, **52**(1), 57-82.

Cheng, Y., M. H. Liu and J. Qian (2006). "Buy-side analysts, sell-side analysts, and investment decisions of money managers." Journal of financial and quantitative analysis **41**(1): 51-83.

Cialdini, R. B. (2009). Influence: science and practice. Harlow, Pearson Education.

Cici, G. and C. Rosenfeld (2016). "A study of analyst-run mutual funds: The abilities and roles of buy-side analysts." Journal of empirical finance **36**: 8-29.

Clarke, J. E., A. Khorana, A. Patel and P. R. Rau (2011). "Independents' day? Analyst behavior surrounding the Global Settlement." Annals of finance **7**(4): 529-547.

Coval, J. D. and T. J. Moskowitz (2001). "The geography of investment: Informed trading and asset prices." Journal of political economy **109**(4): 811-841.

Cowen, A., B. Groysberg and P. Healy (2006). "Which types of analyst firms are more optimistic?" Journal of accounting and economics **41**(1-2): 119-146.

Crawford, S., W. Gray, B. R. Johnson and R. A. Price III (2017). "What motivates buy-side analysts to share recommendations online?" Management science **64**(6): 2574-2589.

Da, Z., K. P. Hong and S. Lee (2016). "What drives target price forecasts and their investment value?" Journal of business finance and accounting **43**(3-4): 487-510.

De Prado, M. L. (2018). Advances in financial machine learning. Hoboken, John Wiley & Sons.

Dechow, P. M. and H. You (2017). "Determinants of errors in analysts' target price implied returns." Working Paper.

Deloitte (2019). Artificial intelligence: The next frontier for investment management firms.

Demirakos, E. G., N. C. Strong and M. Walker (2010). "Does valuation model choice affect target price accuracy?" European accounting review **19**(1): 35-72.

Ertimur, Y., J. Sunder and S. V. Sunder (2007). "Measure for measure: The relation between forecast accuracy and recommendation profitability of analysts." Journal of accounting research **45**(3): 567-606.

Fabozzi, F. J., S. Focardi and C. Jonas (2010). "Investment management after the global financial crisis." CFA Institute Research Foundation.

Fang, B., Hope, O. K., Huang, Z., & Moldovan, R. (2019). "The Effects of MiFID II on Sell-Side Analysts, Buy-Side Analysts, and Firms". Buy-side analysts and firms

Fama, E. F., & French, S. (1992). "Portfolio performance and the interaction between systematic risk, firm size and price-earnings ratio". Journal of finance, **92**, 654-673.

Fama, E. F. and J. D. MacBeth (1973). "Risk, return, and equilibrium: Empirical tests." Journal of political economy **81**(3): 607-636.

Fehr, E. and S. Gachter (2000). "Cooperation and punishment in public goods experiments." American economic review **90**(4): 980-994.

Frankel, R., & Li, X. (2004). "Characteristics of a firm's information environment and the information asymmetry between insiders and outsiders." Journal of accounting and economics, **37**(2), 229-259.

Frey, S. and P. Herbst (2014). "The influence of buy-side analysts on mutual fund trading." Journal of banking and finance **49**: 442-458.

Garcia, M.-F. (2007). The social construction of a perfect market. Do economists make markets?: on the performativity of economics. F. M. DA MacKenzie, L Siu. Princeton, Princeton University Press.

Gleason, C. A., W. B. Johnson and H. Li (2013). "Valuation model use and the price target performance of sell-side equity analysts." Contemporary accounting research **30**(1): 80-115.

Gordon, P. (2017). Sellside research would be little missed. Financial times 6 February 2017.

Granovetter, M. (1985). "Economic action and social structure: The problem of embeddedness." American journal of sociology **91**(3): 481-510.

Greenwood, R. and D. Scharfstein (2013). "The growth of finance." Journal of economic perspectives **27**(2): 3-28.

Grennan, J. and R. Michaely (2018). "Fintechs and the market for financial analysis." Michael J. Brennan Irish Finance Working Paper Series Research Paper(18-11): 19-10.

Grinblatt, M. and M. Keloharju (2000). "The investment behavior and performance of various investor types: a study of Finland's unique data set." Journal of financial economics **55**(1): 43-67.

Groysberg, B., Healy, P., & Chapman, C. (2008). “Buy-side vs. sell-side analysts’ earnings forecasts”. Financial Analysts Journal, **64**(4), 25-39.

Groysberg, B. and P. M. Healy (2013). Wall Street research: Past, present, and future. Stanford, Stanford University Press.

Groysberg, B., P. M. Healy and D. A. Maber (2011). "What drives sell-side analyst compensation at high-status investment banks?" Journal of accounting research **49**(4): 969-1000.

Gu, S., B. Kelly and D. Xiu (2018). Empirical asset pricing via machine learning, National bureau of economic research.

Guan, Y., C. Li, H. Lu and M. Wong (2018). "Regulations and brain drain: Evidence from wall street star analysts' career choices." Management science Forthcoming.

Haig, A. and W. Rees (2015). "Paying for research." Journal of the CISI.

Haig, A. and W. Rees (2016). "The changing market for investment research: The impact of innovations in the payment mechanism on market structure and information flows." Working Paper.

Hashim, N. and N. Strong (2015). "Do formal risk assessments improve analysts' target price accuracy?" Working Paper.

Healy, P. M. and K. G. Palepu (2003). "The fall of Enron." Journal of economic perspectives **17**(2): 3-26.

Hendry, J., P. Sanderson, R. Barker and J. Roberts (2016). "Owners or traders? Conceptualizations of institutional investors and their relationship with corporate managers." Human relations **59**(8): 1101-1132.

Ho, T. Q., N. Strong and M. Walker (2018). "Modelling analysts' target price revisions following good and bad news?" Accounting and business research **48**(1): 37-61.

Imam, S., R. Barker and C. Clubb (2008). "The use of valuation models by UK investment analysts." European accounting review **17**(3): 503-535.

Imam, S. and C. Spence (2016). "Context, not predictions: a field study of financial analysts." Accounting, auditing and accountability journal **29**(2): 226-247.

Ince, O. S. and R. B. Porter (2006). "Individual equity return data from Thomson Datastream: Handle with care!" Journal of financial research **29**(4): 463-479.

Jackson, A. R. (2005). "Trade generation, reputation, and sell-side analysts." The journal of finance **60**(2): 673-717.

Jacob, J., S. Rock and D. P. Weber (2008). "Do non-investment bank analysts make better earnings forecasts?" Journal of accounting, auditing and finance **23**(1): 23-61.

Joos, P., J. D. Piotroski and S. Srinivasan (2016). "Can analysts assess fundamental risk and valuation uncertainty? An empirical analysis of scenario-based value estimates." Journal of financial economics **121**(3): 645-663.

Joos, P. R. and J. D. Piotroski (2017). "The best of all possible worlds: unraveling target price optimism using analysts' scenario-based valuations." Review of accounting studies **22**(4): 1492-1540.

Kadan, O., L. Madureira, R. Wang and T. Zach (2008). "Conflicts of interest and stock recommendations: The effects of the global settlement and related regulations." The review of financial studies **22**(10): 4189-4217.

Kahn, R. (2018). The Future of Investment Management, Research Foundation of the CFA Institute.

Kang, Q., X. Li and T. Su (2018). "Sell-side financial analysts and the CFA® program." Financial analysts journal **74**(2): 70-83.

Kay, J. (2015). Other people's money: Masters of the universe or servants of the people? London, Profile Books.

Kothari, S., E. So and R. Verdi (2016). "Analysts’ forecasts and asset pricing: A survey." Annual review of financial economics **8**: 197-219.

Kryzanowski, L., M. Galler, and D. W. Wright. (1993) "Using artificial neural networks to pick stocks." Financial Analysts Journal **49**, no. 4: 21-27

Li, X. and R. N. Sullivan (2011). "A dynamic future for active quant investing." The journal of portfolio management **37**(3): 29-36.

Lui, D., S. Markov and A. Tamayo (2007). "What makes a stock risky? Evidence from sell-side analysts' risk ratings." Journal of accounting research **45**(3): 629-665.

Lui, D., S. Markov and A. Tamayo (2012). "Equity analysts and the market's assessment of risk." Journal of accounting research **50**(5): 1287-1317.

Maber, D. A., B. Groysberg and P. M. Healy (2014). "The use of broker votes to reward brokerage firms' and their analysts' research activities." Working paper.

MacKenzie, D. (2008). An engine, not a camera: How financial models shape markets. Cambridge, MIT Press.

MacKenzie, D. A., F. Muniesa and L. Siu (2007). Do economists make markets?: on the performativity of economics, Princeton University Press.

Malloy, C. J. (2005). "The geography of equity analysis." The journal of finance **60**(2): 719-755.

Mauss, M. (2002). The gift: The form and reason for exchange in archaic societies. London, Routledge.

Meehl, P. E. (1957). "When shall we use our heads instead of the formula?" Journal of counseling psychology **4**(4): 268.

Mehran, H. and R. M. Stulz (2007). "The economics of conflicts of interest in financial institutions." Journal of financial economics **85**(2): 267-296.

Miller, E. M. (1977). "Risk, uncertainty, and divergence of opinion." The journal of finance **32**(4): 1151-1168.

Mohanram, P. S. (2005). "Separating winners from losers among lowbook-to-market stocks using financial statement analysis." Review of accounting studies **10**(2-3): 133-170.

Newey, W. K. and K. D. West (1987). "Hypothesis testing with efficient method of moments estimation." International economic review: 777-787.

Offer, A. (1997). "Between the gift and the market: the economy of regard." The economic history review **50**(3): 450-476.

Ou, J. A. and S. H. Penman (1989). "Financial statement analysis and the prediction of stock returns." Journal of accounting and economics **11**(4): 295-329.

Petersen, M. A. (2009). "Estimating standard errors in finance panel data sets: Comparing approaches." The Review of financial studies **22**(1): 435-480.

Pinto, J. E., T. R. Robinson and J. D. Stowe (2019). "Equity valuation: A survey of professional practice." Review of financial economics **37**(2): 219-233.

Piotroski, J. D. (2000). "Value investing: The use of historical financial statement information to separate winners from losers". Journal of Accounting Research, **38**, 1-41.

Poon, S.-H. and C. W. Granger (2005). "Practical issues in forecasting volatility." Financial analysts journal: 45-56.

Poon, S.-H. and C. W. Granger (2003). "Forecasting volatility in financial markets: A review." Journal of economic literature **41**(2): 478-539.

Porta, R. L., F. Lopez-de-Silanes, A. Shleifer and R. W. Vishny (1998). "Law and finance." Journal of political economy **106**(6): 1113-1155.

SEC (2019). SEC Announces Extension of Temporary Measure to Facilitate Cross-Border Implementation of the European Union's MiFID II's Research Provisions. SEC Press Release https://www.sec.gov/news/press-release/2019-229

Shane, P. B., S. Rock and S. Ramnath (2006). "Financial analysts' forecasts and stock recommendations: A review of the research." Foundations and trends in finance **2**(4): 311-421.

So, E. and C. M. C. Lee (2015). "Alphanomics: the informational underpinnings of market efficiency." Foundations and trends in accounting **9**(2-3): 59-258.

Spence, C., M. Aleksanyan, Y. Millo, S. Ima and S. Abhayawansa. (2019). "Earning the 'write to speak': sell-side analysts and their struggle to be heard." Contemporary accounting research **Forthcoming**.

Swales Jr, G. S., and Y. Yoon. (1992). Applying artificial neural networks to investment analysis. Financial Analysts Journal, **48**(5), 78-80.

Szado, E., H. Kazemi and T. Schneeweis (2018). "Option informed stock picking." The journal of alternative investments **21**(1): 48-66.

Tetlock, P. (2005). Expert political judgment: How good is it? how can we know? Princeton, Princeton University Press.

Wahlen, J. M. and M. M. Wieland (2011). "Can financial statement analysis beat consensus analysts' recommendations?" Review of accounting studies **16**(1): 89-115.

Walker, O. and C. Flood (2017). MiFIDII leads to exodus of sellside analysts. Financial times 22 June 2018.

White, H. C. (1981). "Where do markets come from?" American journal of sociology **87**(3): 517-547.

Womack, K. L. (1996). "Do brokerage analysts' recommendations have investment value?" The journal of finance **51**(1): 137-167.

Wu, M., M. Wilson, and Y. Wu. (2017). Was the Global Settlement effective in mitigating systematic bias in affiliated analyst recommendations? Journal of business ethics, **146**(3), 485-503.